

AD-A237 105



ETL-0584

An Automated Software
System For Updating Digital
Terrain Databases From
All-Source Imagery, Phase I
SBIR

J. Curlander
W. Kober
J. Thomas

Vexcel Corporation
2477 55th Street
Boulder, CO 80301

A. Stocker

Space Computer Corporation
2800 Olympic Boulevard
Santa Monica, CA 90404-4119

February 1991

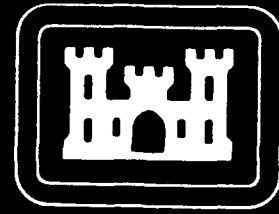
Approved for public release; distribution is unlimited.

Prepared for:

U.S. Army Corps of Engineers
Engineer Topographic Laboratories
Fort Belvoir, Virginia 22060-5546

91 6 19 011

91-02596



E

T

L



Destroy this report when no longer needed.
Do not return it to the originator.

The findings in this report are not to be construed as an official
Department of the Army position unless so designated by other
authorized documents.

The citation in this report of trade names of commercially available
products does not constitute official endorsement or approval of the
use of such products.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE February 1991	3. REPORT TYPE AND DATES COVERED Contract Report	
4. TITLE AND SUBTITLE AN AUTOMATED SOFTWARE SYSTEM FOR UPDATING DIGITAL TERRAIN DATABASES FROM ALL-SOURCE IMAGERY, Phase I SBIR			5. FUNDING NUMBERS DACA76-90-C-0013	
6. AUTHOR(S) J. Curlander W. Kober J. Thomas A. Stocker				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Vexcel Corporation 2477 55th Street Boulder, CO 80301 Space Computer Corp. 2800 Olympic Blvd. Santa Monica, CA 90404-4119			8. PERFORMING ORGANIZATION REPORT NUMBER ETL-0584	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Engineer Topographic Laboratories Fort Belvoir, Virginia 22060-5546			10. SPONSORING/MONITORING AGENCY REPORT NUMBER ETL-0584	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) VEXCEL Corporation's Phase I SBIR research effort for ETL concentrated on the feasibility of creating the primary tools for the prototype development in Phase II of a digital change detection work station. This system is intended to be capable of detecting long-term (6 months to 1 year) and/or seasonal changes from all-source imagery. The system is intended to be hosted on a SUN-4 platform operating under a UNIX/C software environment. The emphasis of the present Phase I effort was on the two major technical challenges for the development of such a system: precision image registration and robust change detection and analysis. Most of this effort was directed toward automated SAR-optical image registration and automated change cueing experiments. Change cueing is an initial step in change detection for identifying regions where possible change events may have occurred. The automated registration and cueing efforts were successful over the data sets tested. These data sets did not contain appreciable terrain-induced distortions. Theoretical improvements for algorithms are recommended for addressing such terrain-induced registration complications in Phase II. Recommendations are also made for the inclusion of interactive capabilities to supplement the automated registration techniques.				
14. SUBJECT TERMS			15. NUMBER OF PAGES 122	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT	

Table of Contents

- 0. Executive Summary
- 1. Introduction to Phase I Report
 - 1.1 Background and Objectives of Phase I Research
 - 1.2 Summary of Conclusions
 - 1.3 Organization of the Report
- 2. Pre-Processing
 - 2.1 Characterization of Imagery Set
 - 2.2 Optical Images
 - 2.3 SAR Images
 - 2.4 Infrared Images
 - 2.5 Pre-Processing Results
- 3. Approximate Registration of SAR-Optical Imagery
 - 3.1 General Issues and Strategies
 - 3.2 Methods
 - 3.2.1 Area-Based
 - 3.2.2 Contour-Based
 - 3.2.3 Combined
 - 3.3 Results
 - 3.4 Remaining Problems
- 4. Sub-Pixel Registration Estimation
 - 4.1 Estimation of Local Translation
 - 4.1.1 Algorithm #1
 - 4.1.2 Algorithm #2
 - 4.2 Global Interpolation of Local Estimates
 - 4.3 Image Resampling

[illegible]

- 4.4 Results
- 4.5 Remaining Problems
- 5. Potential Change Cueing
 - 5.1 Methods for Change Cueing
 - 5.2 Applications to Multi-Sensor Image Sets
 - 5.3 Results
 - 5.4 Remaining Problems
- 6. Preliminary Analysis of Effects of Variations in Imaging Scenarios
 - 6.1 Variable SAR Scenario Effects
 - 6.2 Variable Optical Scenario Effects
- 7. Conclusions
 - 7.1 Summary of Results and Remaining Problems
 - 7.2 Recommendations for Phase II
- 8. References
- 9. Appendix
 - 9.1 Additional Imagery
 - 9.2 Polarimetry Definitions
 - 9.3 Resampling Kernels
 - 9.4 LOWTRAN7 Atmospheric Model

0. Executive Summary

This Phase I SBIR research effort concentrated on the feasibility of creating the primary tools for the prototype development in Phase II of a digital change detection workstation. This system is intended to be capable of detecting long-term (6 months to one year) and/or seasonal changes from all-source imagery. The system is intended to be hosted on a SUN-4 platform operating under a UNIX/C software environment.

The emphasis of the present effort was on the two major technical challenges for the development of such a system: precision image registration and robust change detection and analysis. At the direction of the ETL customer, most of this effort was directed toward automated SAR-optical image registration and automated change cueing experiments. Change cueing is an initial step in change detection for identifying regions where possible change events may have occurred.

The automated registration effort was successful over the data sets tested. These data sets did not contain appreciable terrain-induced distortions. Additional testing and refinement of the present algorithms is recommended for Phase II.

Some theoretical developments for algorithms are made for addressing such terrain-induced complications in Phase II. The implementation, testing, and modifications of such algorithms will also be a Phase II priority.

However, a useful system must always allow for interactive intervention by a human operator. This is because automated techniques are not guaranteed to perform perfectly 100% of the time. Therefore, recommendations are also made for an interactive system to supplement the automated registration techniques.

The second main effort concerned change cueing. This effort was also successful over the data set which was attempted. However, more testing must be performed in Phase II over additional imagery sets.

Change analysis is the subsequent step which evaluates such cued regions as being legitimate changes or not. Change analysis was not pursued during this Phase I effort, and will only be pursued to a limited degree in Phase II. This is because the difficulty of this problem lends itself better to an interactive approach with a human operator. However, the operator is expected to be selectively cued to only a small number of possible change events.

An initial effort was also included for describing some of the image variabilities that occur as the the sensing scenarios change.

Following a summary in section 7.1 of results obtained so far, an initial summary of Phase II requirements appears in section 7.2.

1. Introduction to Phase I Report

The Phase I Technical Objectives are summarized in section 1.1, an overview of the conclusions appears in 1.2, and the layout of this report is described in section 1.3.

1.1 Background and Objectives of Phase I Research

The two technical objectives concerned the technically demanding tasks of precision registration and change detection and analysis.

Technical Objective #1: Demonstrate the feasibility of precision registration of multi-source imagery, both image-image as well as image-map.

Technical Objective #2: Demonstrate the feasibility of robust change detection and analysis for multi-source imagery.

The Phase I effort therefore concentrated on registration experiments and the feasibility of initial cueing for change detection. The registration effort examined both rough registration as well as sub-pixel estimation for SAR-optical pairs of imagery.

Change detection involves cueing potential changes and evaluating these cued regions as legitimate changes or as image artifacts due to other causes. The Phase I research only dealt with cueing potential change events. The change events examined were actual changes over time and synthetic changes inserted into imagery.

The examination of these issues was crucial for achieving the ultimate Phase II goal of creating a workstation for automatic change detection of all-source imagery.

1.2 Summary of Conclusions

The image-image registration efforts were encouraging. The two types of rough registration methods were area-based and contour-based. These algorithms operated under the assumption of little terrain distortion to the imagery. A contour-based algorithm for dealing with the problem of registering SAR image pairs containing considerable terrain relief is outlined. Fuller development and testing of this procedure is a Phase II issue.

A two-stage method of combining the area-based and contour-based algorithms showed some improvement over the performance of each separately. In particular, small residual amounts of rotation on the order of a degree were removed by this procedure.

Two sub-pixel registration estimation algorithms are also presented, with extensive testing results so far available for one of them. The further investigation of the other algorithm will be undertaken in Phase II.

The following statistics summarize registration accuracies achieved in testing:

- o area-based
 - K-L algorithm (see section 3.2.1.1): ± 1 pixel in 50% of cases,
 - MNF algorithm (see section 3.2.1.2): ± 1 pixel in 75% of cases,
- o contour-based: ≤ 2 pixels in all cases,
- o combined area-contour: ± 1 pixel (using MNF area method)
- o sub-pixel registration:
 - Algorithm #1 (see section 4.1.1): $\sim 5 \times 10^{-4}$ pixel
(using lower frequencies, and without spectral leakage filtering)
 - Algorithm #2 (see section 4.1.2): $< 10^{-3}$ pixel (best results using Hanning filter for filtering spectral leakage)

The sub-pixel results were obtained on simulated data.

The other main technical area investigated was change cueing on the pixel level, given a registered image pair. Both actual changes and simulated changes were examined. The results of testing revealed excellent cueing even for small (two pixels) targets as long as the local image-image correlation of the background was high, with the performance degrading as this correlation decreases.

From a signal processing standpoint, this performance dependence on the background correlation levels is unavoidable for pixel level processing. However, considerably more progress can be achieved using processing methods which effectively increase these background correlation levels by restricting attention to selective frequency regions.

Progress beyond what can be achieved using such enhanced pixel-level processing would probably require higher-level procedures on the object-level. Such methods will be required to make hypotheses on the existence of objects based on pattern analysis, as opposed to simply thresholding based on local statistics.

Despite the encouraging results using the automated methods, it is strongly recommended that the Phase II workstation contain capabilities for interactive as well as automated modes of operation. Such a dual capability allows the use of a human operator to examine and assess the results generated by automated registration procedures, make corrections if needed, and to provide initial offsets for difficult or ambiguous cases.

For change detection, the use of automated change cueing requires only selective attention by the operator, but uses his/her superior judgement for evaluating cues as legitimate targets.

1.3 Organization of the Report

The pre-processing methods appear in section 2, initial rough and sub-pixel registration methods are in sections 3 and 4, and change cueing is discussed in section 5. Some observations on the variability of these procedures for changing imaging scenarios is presented in section 6.

Finally, some conclusions on the practicality of these methods are discussed in section 7.

2. Pre-Processing

2.1 Characterization of Imagery Set

Any study of registration and change detection methods for multi-source, multi-spectral imagery first requires compiling a representative set of imagery containing the relevant objects of interest. More precisely for change detection, temporal data sets depicting the same region at different time periods are what is needed. Additionally, multi-source or multi-spectral temporal data sets provide opportunities for registration and change detection experiments between imagery types.

During the present Phase I research, the effort concentrated on matching and change detection between synthetic aperture radar (SAR) and electro-optical (EO) Landsat TM imagery.

One multi-source data set consists of the JPL quadpole aircraft SAR of the Raisin City CA site and collateral coverage from 7 bands of Landsat TM. The SAR and TM data represent imaging times. Also, SIR-B imagery of the same site was available.

A second data set consists of SEASAT SAR and 7 bands of TM data of the Yuma AZ site. Again, the SAR and TM data were imaged at different times.

The Landsat TM data was resampled to match the resolution of the SEASAT data for registration experiments.

Table 2-1 Imagery set

Sensor	Region	Resolution (range x azimuth)	Wavelength	Polarization	Flying Altitude
JPL quad-pole SAR	Raisin City, CA	7.49m x 10.58m	L-Band	complex quadpole	12.22 km
SIR-B SAR	Raisin City, CA	34.6m x 28.5m	L-Band 23 cm	H-H	235 km
SEASAT SAR	Yuma, AZ	25m x 25m	L-Band 24 cm	H-H	800 km
Landsat-4 TM	Raisin City, CA Yuma, AZ	30m x 30m*	band 1 .45 μ m - .52 μ m band 2 .52 μ m - .60 μ m band 3 .63 μ m - .69 μ m band 4 .76 μ m - .90 μ m band 5 1.55 μ m - 1.75 μ m band 6 10.4 μ m - 12.5 μ m band 7 2.08 μ m - 2.35 μ m	N/A	705 km

* band 7 Landsat resolution 120m

Table 2-1: Resolutions for Sensor Types

The resolution of the various data sets is given in Table 2-1. Examples of the relevant imagery appear in Fig. 2-1 to 2-5.

The JPL quadpole data set involves independent measurements for the horizontal and vertical polarized SAR returns. From these data, one can calculate the radar returns arising from any hypothetically transmitted linearly polarized signals. Recall that the electric (E) field of an electromagnetic (EM) wave instantaneously resides in a plane of vibration. If this plane of vibration is constant we say that the EM wave is linearly polarized.

The most general case is that this vibration plane rotates, with the magnitude of the E-field vector tracing an ellipse when projected into a plane perpendicular to the propagation vector. In this case we say the EM wave is elliptically polarized. A special case of elliptic polarization is circular polarization.

Elliptically polarized waves can be considered as the superposition of two orthogonal linearly polarized waves with a phase difference. Therefore, many, though not all, analyses of polarization involve only the linear case.

A formalism for representing the outcome of scattering coherent, polarized waves is by the use of the scattering, or Jones, matrix [s]. When post-multiplied by the incident electric field, or Jones vector, the resultant scattered electric field is determined, ie:

$$\begin{bmatrix} E_x^{(s)}(t) \\ E_y^{(s)}(t) \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \begin{bmatrix} E_x^{(t)}(t) \\ E_y^{(t)}(t) \end{bmatrix}$$

Another formalism for representing incoherent scattering involves the four component Stokes vector (S) and the (4x4) Mueller (also called phase or Stokes) matrix [F]. Operationally, incoherent scattering is described in an ensemble sense using matrix multiplication:

$$\begin{bmatrix} S_0^s \\ S_1^s \\ S_2^s \\ S_3^s \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{33} & F_{34} \\ F_{41} & F_{42} & F_{43} & F_{44} \end{bmatrix} \begin{bmatrix} S_0^t \\ S_1^t \\ S_2^t \\ S_3^t \end{bmatrix}$$

Backscatter from terrain involves depolarization and incoherent returns. Instantaneously, loss of polarization does not exist, ie. it is inherently a time-averaged concept. Because the scattering matrix formalism always involves full polarization and complete coherence, it is an instantaneous concept and cannot be used in an ensemble sense. For this reason and because of noise reduction, most polarimetry data is in the form of the Mueller matrix.

As described in [Zebker et al,87], the JPL imaging radar polarimeter measures both intensity and relative phase of radar backscatter as a function of transmitted and received polarizations. This was accomplished by adding a dual-polarized antenna and a four channel data system to the JPL aircraft SAR.

The amplitudes and phases of all elements of the scattering matrix are therefore measured for each pixel in the SAR image. This allows the synthesis of any combination of transmitted and received antenna polarizations. However, as pointed out in [Lukert and Blanchard,88], such accurate depolarization measurements require adequate antenna polarization isolation. Otherwise like-polarization feed-through can exist in the depolarized channel, and this effect can couple with PRF-related azimuth ambiguities to affect pixels other than the one being processed.

Another consideration is the use of the Mueller matrix rather than the scattering matrix. Because of statistical variations caused by noise, spatial averaging of radar power measurements is used at a cost of reduced resolution. In such a case, the assumption of uncorrelated phases of EM waves from individual scatters allows the (incoherent) Stokes vector of the sum of these waves to be the sum of their individual Stokes vectors. This allows the summation of their individual Mueller matrices to form a composite Mueller matrix for the composite scatter composed of individual scatters.

For a fuller discussion of polarization concepts and definitions, see Appendix 9.2.



Fig. 2-1: Landsat Band 4, Yuma AZ Site



Fig. 2-2: SEASAT, Yuma AZ Site

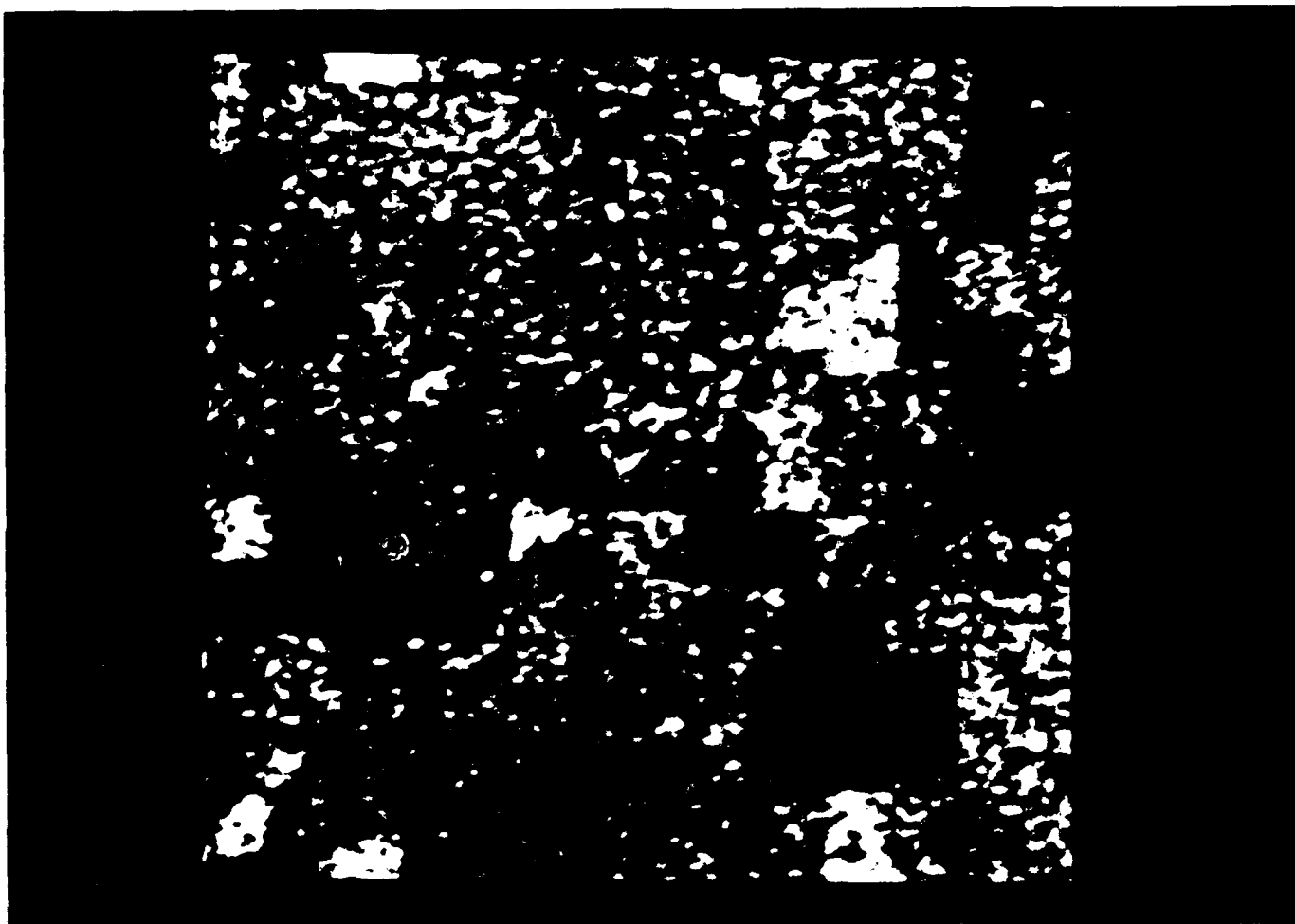


Fig. 2-3: SIR-B SAR, Raisn City CA Site

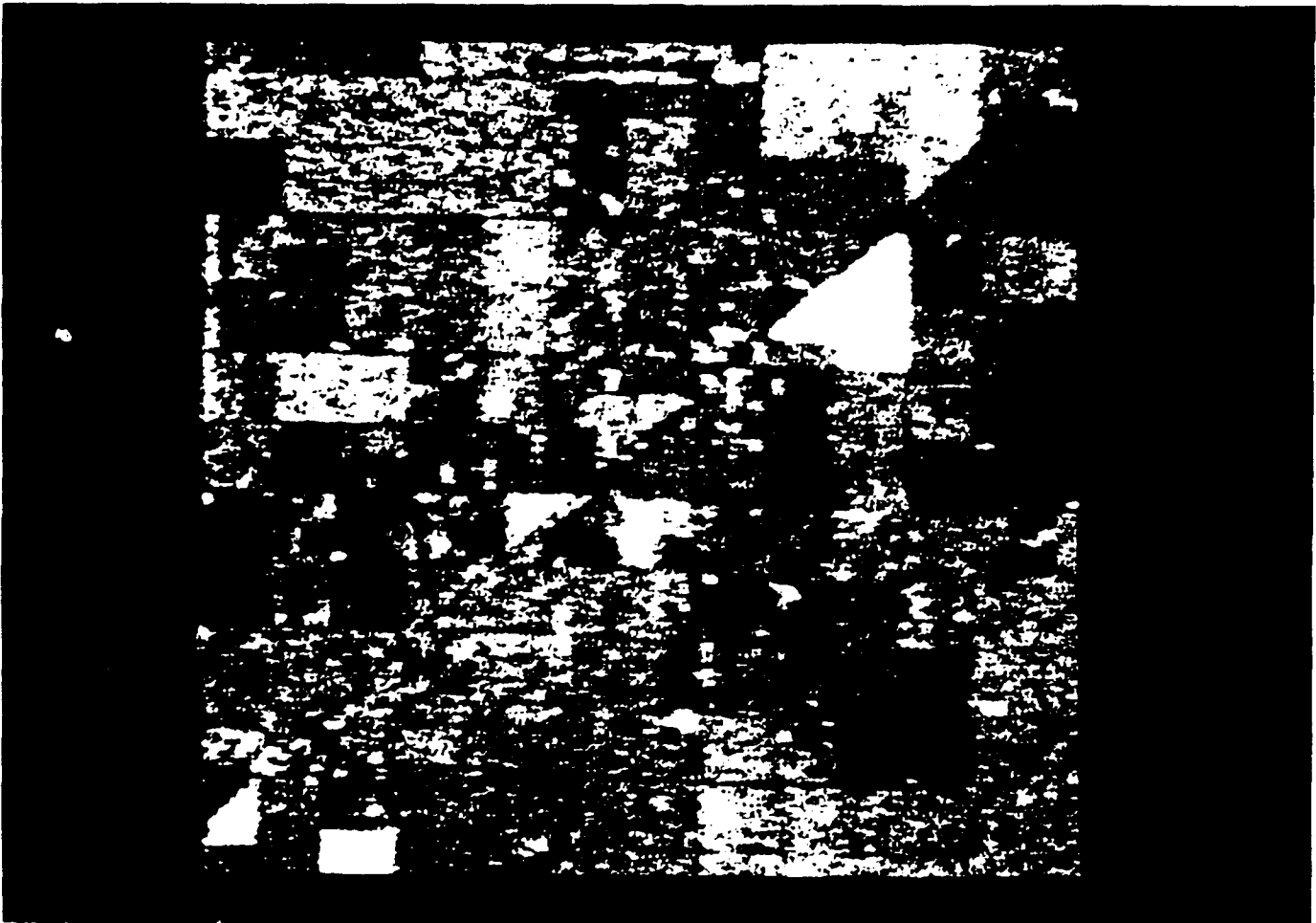


Fig. 2-4: JPL Aircraft SAR (H-H), Raisin City CA Site

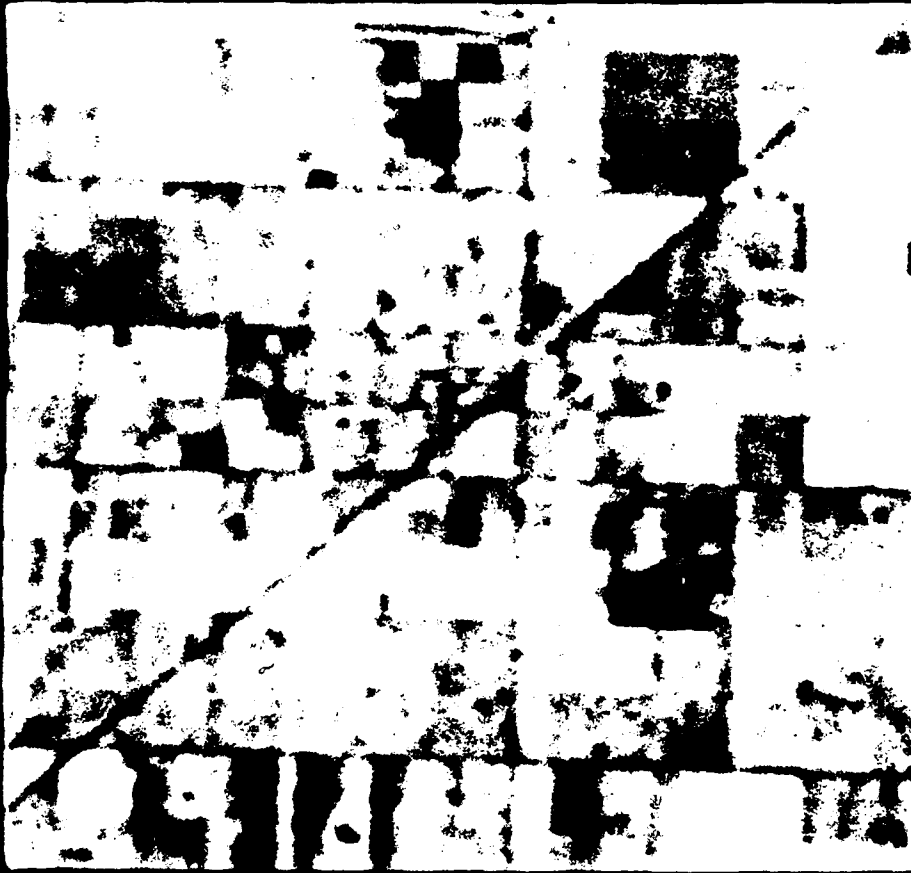


Fig. 2-5: Landsat Band 4, Raisin City CA Site

2.2 Optical Images

Several image pre-processing algorithms suitable for optical imagery were investigated. These included procedures for enhancing edges, reducing noise, and increasing the visual similarity of the optical and SAR images to be matched.

The strategy for matching SAR to optical imagery was to preferentially use the lower noise optical images to cue features for matching in the SAR images. Certain features, particularly those with long linear signatures, can often be segmented from optical imagery more easily and directly than from SAR data. Several pre-processing schemes were tested with this idea in mind.

For edge enhancement, a method which uses the sigma filter [Lee,83] was found to yield well-defined edge structures. The sigma filter is based on the use of a Gaussian distribution model for the pixel intensity values in a local neighborhood. The assumption is that only those pixels whose intensity values lie beyond two standard deviations (2σ) from the neighborhood sample mean comprise another population. This population beyond $2-\sigma$ represents either an edge structure or shot noise, whereas the population within $2-\sigma$ represents an unchanging intensity field, generally corrupted by speckle noise.

Therefore only pixels within the $2-\sigma$ region are averaged to reduce the effects of speckle noise, resulting in an intensity-smoothed pixel. An ad-hoc thresholding procedure is used to infer the presence of shot noise, followed by smoothing of those pixels involving only their 4 nearest neighbors.

The present edge detection procedure uses the sigma filtered image and the original image together. The pixel-by-pixel differences of the two images produces a well-defined edge map with stronger edges having brighter intensities than the edges in the original image.

This enhancement occurs because of the smoothing of flat regions along with a simultaneous identity operation applied to the edges. The image difference operation thus results in amplification immediately next to those areas having abrupt changes and are therefore changed by the sigma filter, whereas flatter regions are suppressed because of their similarity in both images.

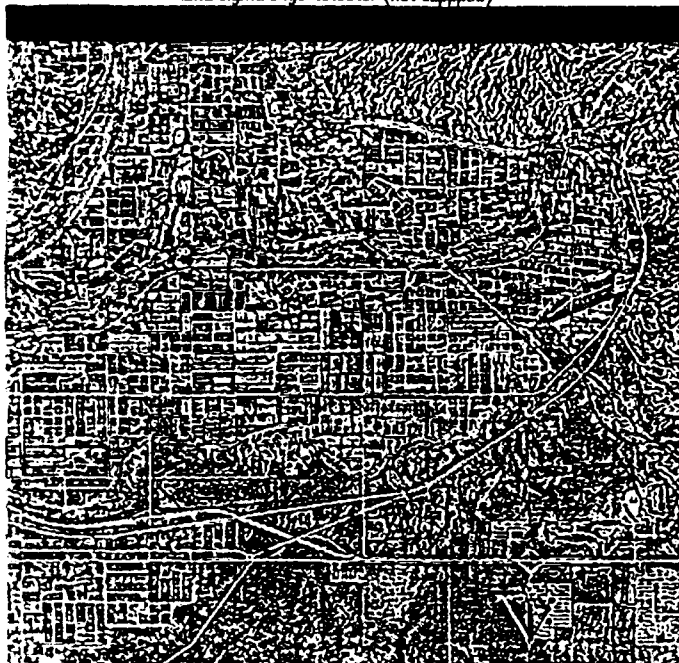
Note that the original edges themselves are not modified by the sigma filter, and thus are also suppressed in the difference image. Therefore, the resulting edges in the difference image are edges which are slightly displaced on the order of one pixel.

In addition, it was discovered that by allowing the byte values of the image to wrap in the difference of the original image, it was possible to obtain an essentially binary image. This occurred because the edges in the sigma filtered image were slightly brighter than in the original image. As a result of the differencing process, these locations were

assigned a value slightly less than zero while all other locations were assigned a grey value slightly greater than zero. The result was an image whose histogram was dominated by the grey values near zero and near 255, with no pixels having grey values in between.

A sigma filter edge-enhanced Landsat TM band 4 image of the Yuma site and one of the Raisin City site are shown in Fig. 2-6. The effect of the wrapping compared to the clipped difference process can be seen in the images. The grey values of the clipped images have been scaled to enhance the dynamic range of the images.

Yuma sigma edge detector (not clipped)



(a)

Yuma sigma edge detector (clipped)



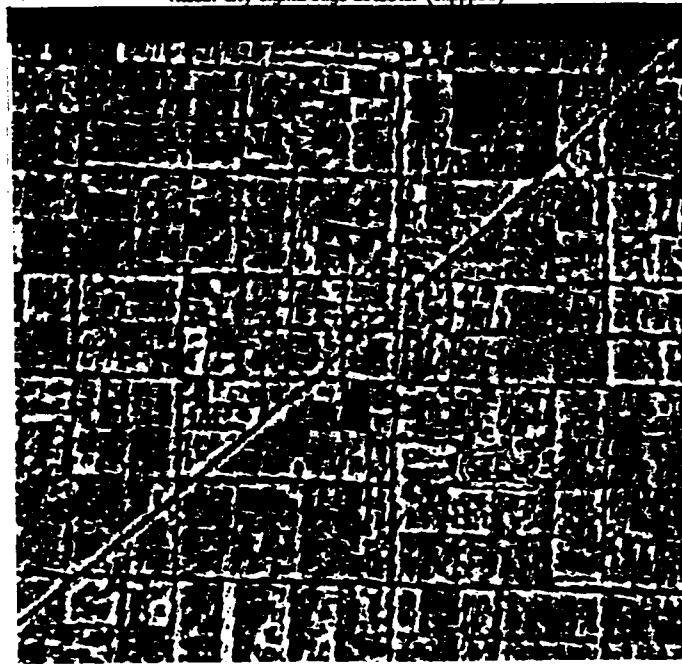
(b)

Raisin City sigma edge detector (not clipped)



(c)

Raisin City sigma edge detector (clipped)



(d)

Figure 2-6 Sigma filter edge enhanced images of Yuma (a) bit-wrapped (b) clipped at zero in the difference, (c-d) similar images of Raisin City.

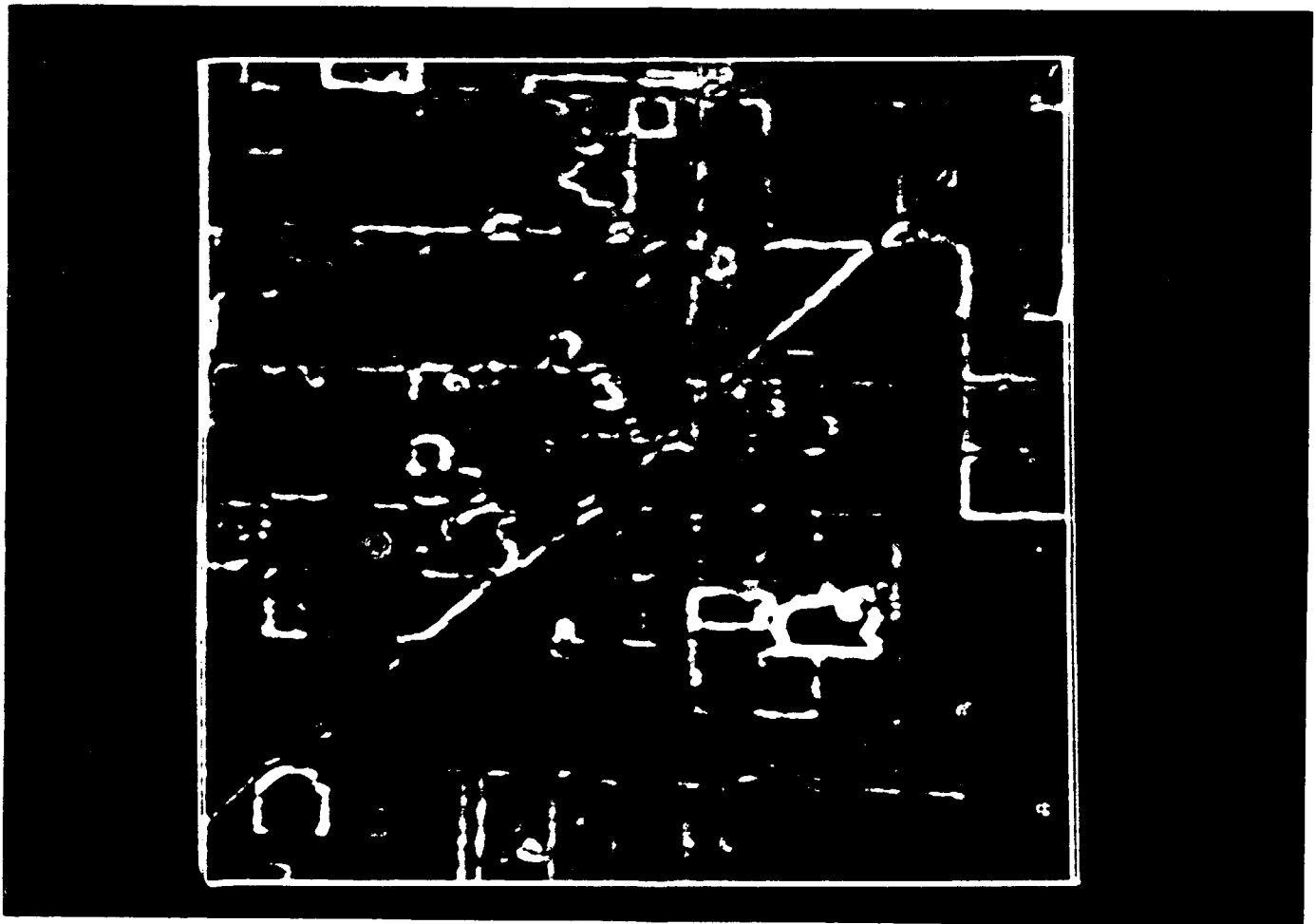


Fig. 2-7: Sobel edge filter enhanced and sigma filter smoothed Landsat Band 4 image of Raisin City site.

The spirit of this method is compatible with the concern voiced in [Adair and Guindon,90] that the usual independent applications of noise suppression and edge enhancement should not be de-coupled. The essential point is that the homogeneity criteria for the adaptive smoothing operator should be consistent with the abruptness detection process in the edge operator.

Another method for edge enhancement, the Sobel operator [Ballard,82], was used effectively as a gradient-type edge detector in bands 1-5 of the Landsat-4 TM data sets. The result of the sobel operator applied to a sigma smoothed Landsat-4 image of the Raisin City site is shown in Fig. 2-7.

Continued effort in Phase II will be concerned with segmentation of edges and contours using such ideas.

2.3 SAR Images

Past work in the field of SAR-optical registration has emphasized the importance of intensity pre-processing the SAR image prior to matching with an optical image. SAR images tend to have lower signal to noise ratios and smaller dynamic ranges than optical images. Another complicating factor is that SAR images involve multiplicative noise, whereas passive optical imagery is corrupted by additive noise.

However, as argued in [Frost et al,82], a more realistic noise model for SAR image processing is multiplicative-convolved noise. This is because the product of the terrain backscatter and coherent fading, ie. the speckle process, is convolved with the radar system's point spread function. Therefore, a straightforward application of the logarithm of the SAR image will not necessarily separate signal and noise into additive components. Direct deconvolution of the signal from this convolved speckle noise process is generally computationally inaccurate because of the low signal to noise ratios common to SAR images. Because of this difficulty, most SAR image processing is simply performed on the basis of a multiplicative noise model even though it is only approximately correct.

Some of the methods for noise reduction in SAR images include locally adaptive signal estimation [Frost et al,82], [Lee,83]. These methods attempt to estimate the signal locally, thereby distinguishing meaningful grey-value transitions from noise-induced changes. Appendix A contains a fuller account of such locally based noise reduction techniques.

Depending on the matching methodology to be used, various nonlinear adaptive spatial filters may be used. For SAR images with a large amount of noise and speckle content, such as is the case with SEASAT and SIR-B imagery as well as to a lesser extent with aircraft imagery, an adaptive filter reduces the effects of noise while still preserving the integrity of edges. This retention of the edge structure is vital to our contour-based matching algorithm.



Fig. 2-8: Comparison of sigma smoothed imagery. Original image upper left, images in upper right, lower left, lower right represent values of smoothing parameter $s = .15, .3, .45$ respectively.

One such filter is the sigma filter, described in section 2.2. VEXCEL has used the sigma filter for filtering SEASAT arctic imagery on another project [McConnell,87]. This filter was used to filter the image prior to successfully thresholding ice floes using histogram techniques.

The sigma filter uses a local statistical model for the flat and edge populations in the SAR image. Increasing the values of sigma lead to images with increasing amounts of smoothing, as shown in Fig. 2-8. Recommendations for sigma values based on maximizing the signal to noise ratio appear in [Lee,83].

Another edge preserving, adaptive noise smoothing algorithm [Nagao et al,79] uses a set of nine masks for each pixel. Each such mask covers a select group of the neighbors of a given pixel. The variance of the pixel intensities of each mask is calculated, and the mean of the pixel mask with the minimum variance is substituted for the original pixel value.

An extension of the sigma filter to quadpole SAR imagery is given in [Lee,90]. The extension of such adaptive smoothing techniques to data sets of polarized SAR imagery represents a promising approach for increasing the signal to noise ratio of the filtered images.

An algorithm based on relaxation [Rosenfeld,82] was attempted for reinforcing "smooth" contours formed by local edges. The algorithm is not rigorous in a probabilistic sense and does not compute updated probabilities in according to Bayes theorem. Nevertheless, it is appealing in a heuristic sense, and seems to correspond to intuitive thinking about locally updating probabilities of an edge being significant based on its immediate surroundings.

This algorithm is briefly given below:

- 1) Estimate initial probabilities $p(i,j)$ based on the likelihood that a pixel is a member of class j . Each pixel in the window of n pixels is pixel i . There are $m=4$ classes $j=1,...,4$, corresponding to the four edge orientations $0^\circ, 45^\circ, 90^\circ, 135^\circ$. The sense of the edges is ignored at this point in this processing for simplicity.
- 2) For each point, determine the compatibility of each edge class probability with that of its eight neighbors. Thus, construct $c(i,j;h,k)$, the compatibility of pixel i with class j , and pixel h with class k . The pixels each have probabilities $p(i,j)$, and $p(h,k)$ of being a member of the specified class. There are eight compatibilities $h=1,...,8$ for each class of pixel i . Therefore, for each pixel there are 32 compatibilities, ie. 4 for each of 8 neighbors.
- 3) Update probabilities by first computing $q(i,j)$, which is the mean of the sum:

$$q(i,j) = (1/n-1) \sum_{h=1}^8 \sum_{k=1}^4 \{c(i,j) p(h,k)\}$$

Then the new updated probability is given by:

$$p_{new}(i,j) = p(i,j)_{old}(1+q(i,j))$$

This algorithm did not work well in its present form for SAR imagery. The results are discussed in section 2.5.

Methods employed for intensity re-mapping include histogram equalization using the cumulative distribution functions of both images [Casteleman,79], histogram re-mapping using the Karhunen-Loeve Transform [Wong,77], and correction using a sensor-dependent, exponential model for the optical image [Wong,80].

2.4 Infrared Images

The NASA airborne Thermal Infrared Multispectral Scanner (TIMS) has six wavelength channels in the long wave infrared (8-12 microns) region of the electromagnetic spectrum. It has an instantaneous field of view of 2.5 mrad, a total field of view of 80 degrees and a nominal ground resolution of 10 meters. Table 2-2 shows the specifications of the six LWIR TIMS bands.

We have acquired two daytime near-nadir TIMS images of Death Valley in the Mojave Desert of California from the Jet Propulsion Laboratory — one taken in June of 1983 and the other in July of 1988. Sortie specifications are shown in Table 2-3. The respective images for band 1 are shown in Figures 2-9(a) and 2-9(b).

<u>Band</u>	<u>Wavelength Limits</u> <u>(microns)</u>	<u>Bandwidth FWHM</u> <u>(microns)</u>
1	8.2 - 8.6	0.4
2	8.6 - 9.0	0.4
3	9.0 - 9.4	0.4
4	9.4 - 10.2	0.8
5	10.2 - 11.2	1.0
6	11.2 - 12.2	1.0

Table 2-2: TIMS Wavebands

	<u>Death Valley 1</u>	<u>Death Valley 2</u>
Sortie Date	6/02/88	7/22/83
Start Time	20:59:53	18:43:38
End Time	20:07:43	18:46:30
Start Latitude	36D 44M 48S N	36D 23M 48S N
End Latitude	36D 0M 0S N	36D 24M 12S N
Start Longitude	117D 1M 24S W	116D 45M 12S W
End Longitude	116D 51M 54S W	117D 1M 24S W

Table 2-3: Specifications for Death Valley Flights

Death Valley is a long narrow valley floor bounded by the Panamint Range on the west and the Funeral and Black mountains on the east, which consist of deformed metamorphic, volcanic, and sedimentary rocks. The floor of Death Valley, which includes the lowest elevation in the United States, is a closed salt pan covered by evaporite deposits — the dark, delineated, stream-like forms near the right side of Figures 2-9(a) and 2-10(b). These intermittent streams on the valley floor are bordered by floodplain deposits of clay and silt. The valley floor is bordered by alluvial fan deposits of gravel eroded from the adjacent mountains — these are the round skirt-like forms to either side of the valley floor. On the Panamint mountains, to the west (left) of the featureless alluvial gravel fans, bedrock is exposed as intrusive bright forms (volcanic rocks) and rough, jagged dark forms (carbonate rocks). Also of interest are two man-made features near the upper right (north-east) of Figure 2-9(b): Furnace Creek Ranch (a mottled square dark feature) and Furnace Creek airstrip (a long thin feature oriented in the north to south direction). The only common man-made feature between these two images is the airstrip.

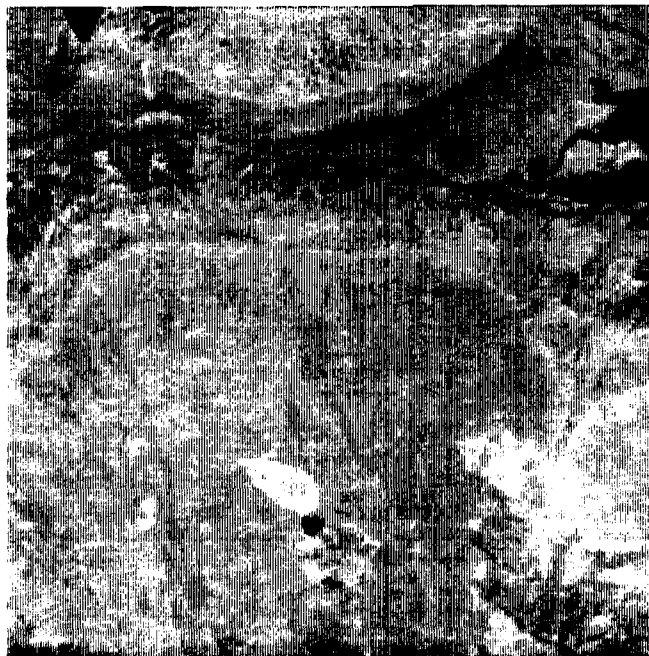
It is evident from Figures 2-9(a) and 2-9(b) that the Death Valley images have different scales and are grossly misregistered. The first step in the preprocessing was therefore to perform a global registration based on a set of 54 manually located control points in the two images. For each image dimension, the measured displacements between corresponding control points were fit to a thin-plate cubic spline and interpolated to generate a warping surface (one for each image dimension). The local 2-D misregistrations provided by these surfaces were then used to resample each of the images using a sliding-window cubic B-spline kernel (discussed in Section 4.3). The globally registered images for TIMS Band 1, shown in Figures 2-10(a) and 2-10(b), visually appear to be quite well aligned. The same global registration step was also applied to the TIMS images in Bands 3 and 5, so that the three bands for each observation could be processed together.

The apparent similarity of multi-band infrared data sets is sometimes greater in certain principle spectral components than in the original bands themselves, due to enhancement of spectrally distinct features in the data. Figures 2-11 through 2-13 compare the three respective principle components computed from Bands 1, 3 and 5 (after global registration) for each of the two Death Valley data sets. The long-term correspondence between certain features in Components 2 and 3 (Fig. 2-12 and 2-13) is particularly striking. Note also that the Furnace Creek airstrip is clearly visible in both second component images as a somewhat crooked dark line near the upper right-hand edge of the frame. Since the airstrip is known to be a linear feature, this shape indicates the presence of residual global registration error near the right frame edge.

2.5 Pre-Processing Results

The SAR imagery used in the present effort contains a large amount of speckle noise. The SIR-B imagery had the smallest signal to noise ratio and was more dominated by speckle than the other SAR types. Moreover, the smaller look angle of the space platform led to a higher degree of geometric distortions.

(a) Death Valley 1 (1983)

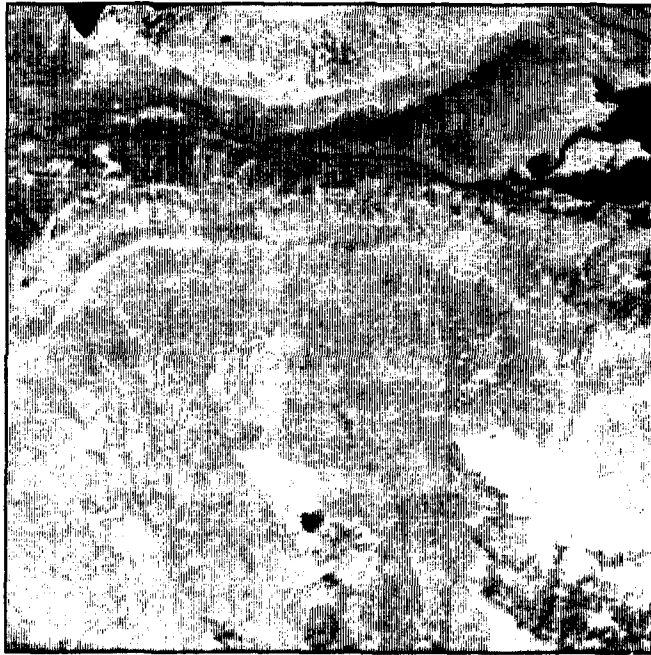


(b) Death Valley 2 (1988)



Fig. 2-9. TlMS Death Valley Thermal Infrared Images (Band 1)

(a) Death Valley 1 (1983)

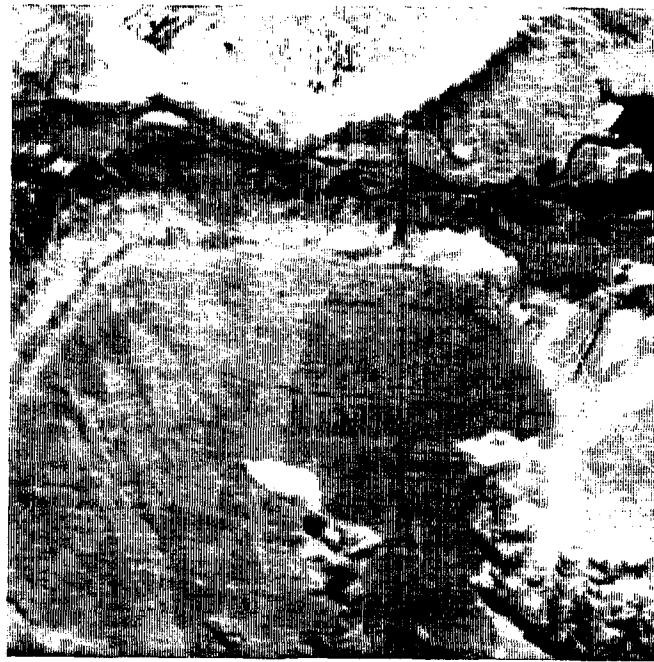


(b) Death Valley 2 (1988)



Fig. 2-10. TMS Death Valley Images after Global Registration (Band 1)

(b) Death Valley 2 (1988)



(a) Death Valley 1 (1983)



Fig. 2-11. First Principle Component Images from TMS Bands 1,3,5

(a) Death Valley 1 (1983)



(b) Death Valley 2 (1988)



Fig. 2-12: Second Principal Component Images from TMS Bands 1,3,5.

(a) Death Valley 1 (1983)



(b) Death Valley 2 (1988)



Fig. 2-13: Third Principal Component Images from TMS Bands 1,3,5.



Fig. 2-14: Sigma filtered SEASAT SAR, smoothing parameter $s=.25$, 5×5 pixel neighborhood.

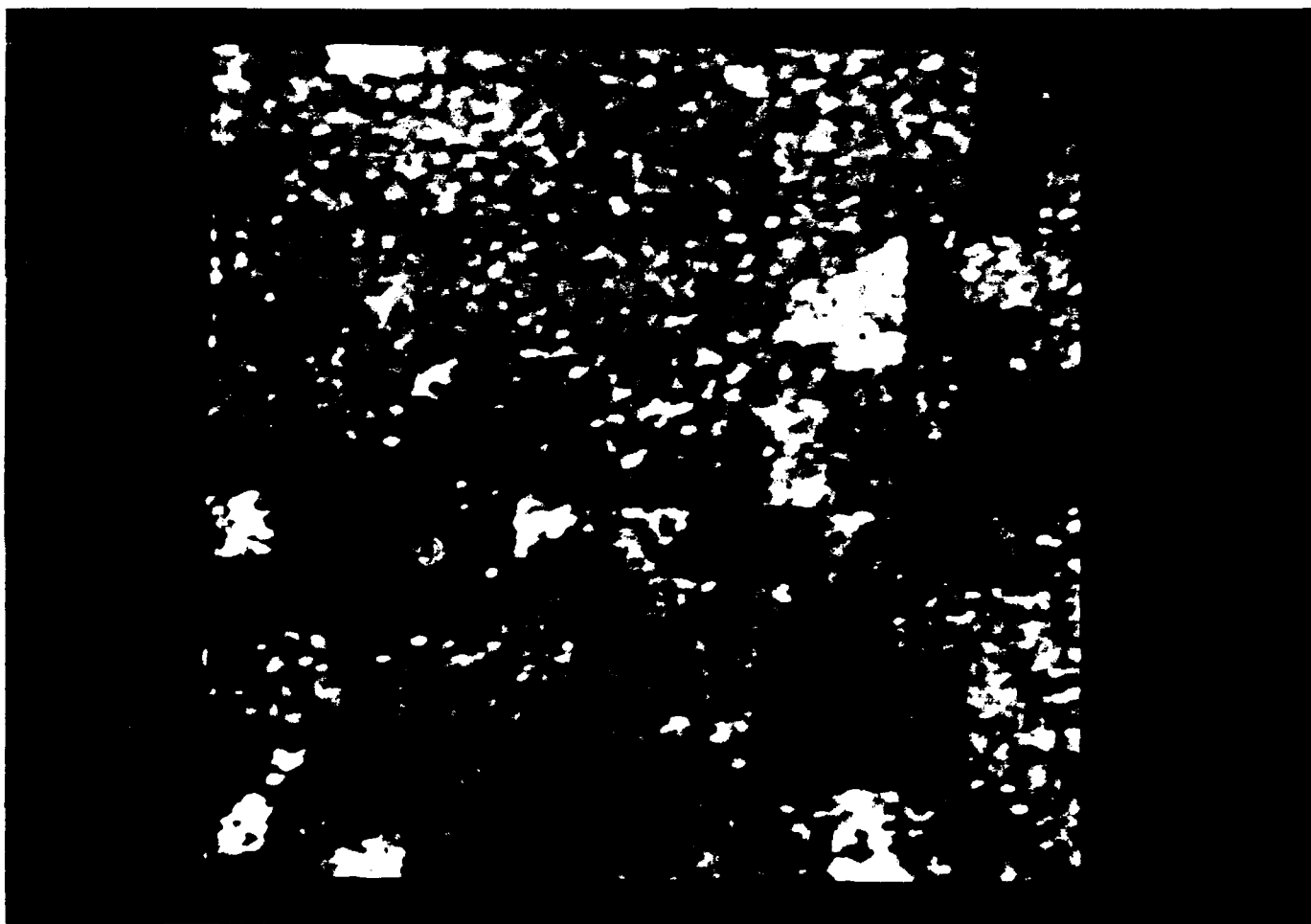


Fig. 2-15: Sigma filtered SIR-B SAR, smoothing parameter $s=.1$, 5x5 pixel neighborhood.



Fig. 2-16: Sigma filtered JPL (H-H) SAR, smoothing parameter $s=.1$, 5×5 neighborhood.

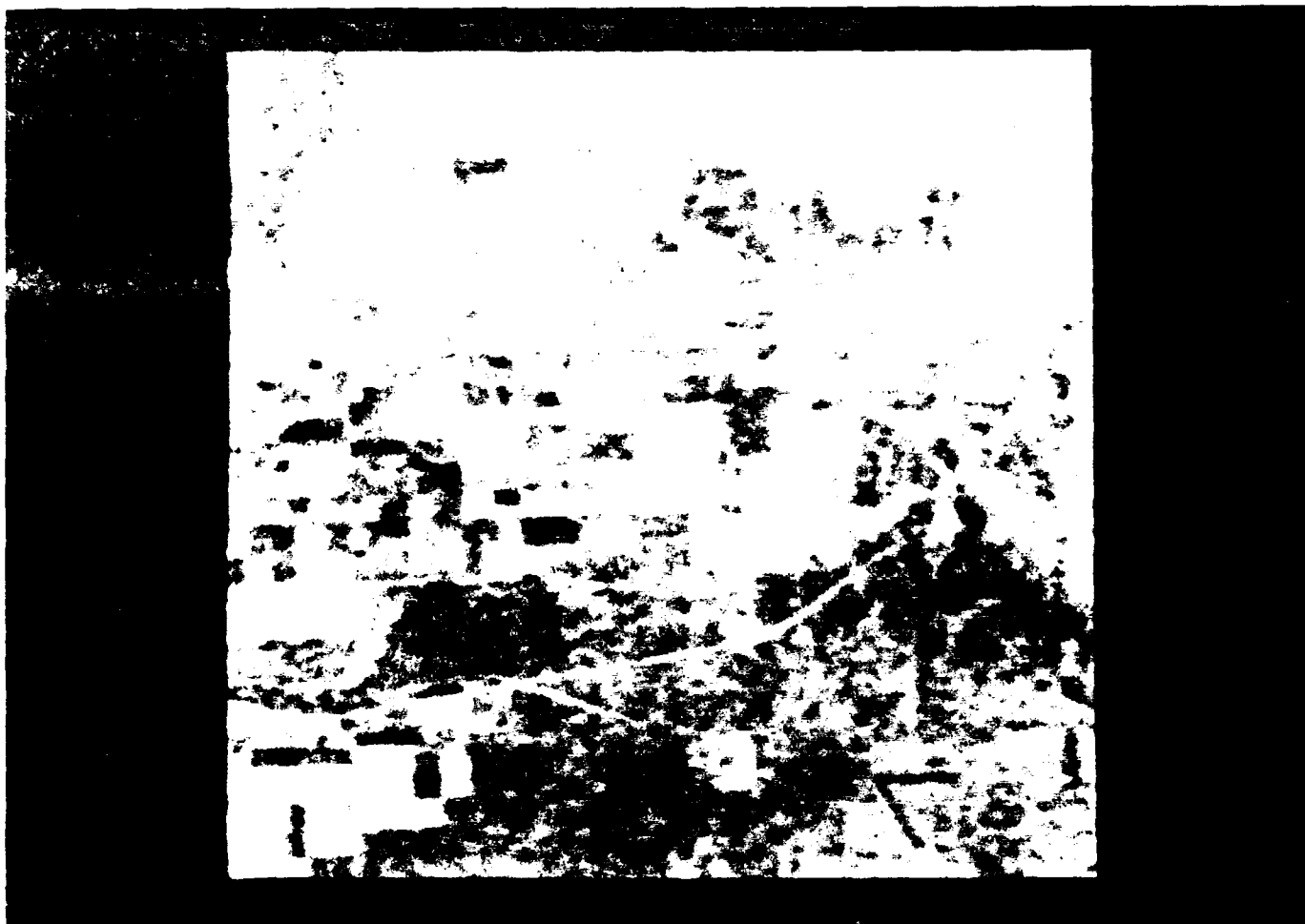


Fig. 2-17: Sigma filtered, histogram remapped SEASAT SAR, smoothing parameter $s = .25$, 5×5 neighborhood.

The sigma filter was used extensively for smoothing the SEASAT, SIR-B, and JPL aircraft quadpole SAR data sets. Several variations on the parameters for filtering were tried. Fig. 2-14 shows a sigma-filtered SEASAT image of the YUMA scene. Fig. 2-15 and 2-16 are sigma-filtered images of the Raisin City for SIR-B and JPL quadpole sensed data.

In order to increase the amenability of the SAR-optical matching process, a global histogram remapping was applied to the SEASAT data based on the histogram of the corresponding optical image. This global filtering step was then followed by applying the locally adaptive sigma filter. The resulting image is shown in Fig. 2-17.

Clearly, these sigma-filtered images show improvements in noise content and the histogram remapping operation improved the visual similarity of the SAR image to the optical (see Fig. 2-1 and 2-2).

Unfortunately, the relaxation procedure for selectively emphasizing long contours [Rosenfeld,82] did not work well with SAR imagery. In particular, most edges were eventually reinforced and strengthened, creating a confusing, cluttered image instead of selectively segmenting only major contours.

It is believed that this problem is because small edge structures that initially get emphasized are allowed by the algorithm to persist as the number of iterations increases. Because of the cluttered nature of the edges in a SAR image, most small edge structures do get emphasized during the initial iterations.

This suggests that a more successful algorithm should be successively taking into account larger neighborhoods that are directionally oriented as the iterations progress, instead of always using a neighborhood of fixed size.

Another modification that may be considered for pre-processing prior to SAR-optical matching is to only reinforce those directions that are contained in a particular streak S that has been extracted in the optical image as part of a feature (the preferential use of streaks extracted first from the optical image is discussed in section 3.2.2.1). Then perform this directionally selective reinforcement successively for all candidate streaks S in the SAR image which are in the region of registration ambiguity for S . Moreover, the compatibilities $C(i,j;h,k)$, discussed in section 2.3, should be computed directly from their neighbors in S using histograms, instead of from the compatibility table presently used.

These modifications will be examined in Phase II.

3. Approximate Registration of SAR-Optical Imagery

Image-image registration methods occur at three scales of accuracy:

- o Very rough registration using geocoding methods, resulting in errors of multiple pixels,
- o Approximate registration using pattern information in the images, resulting in errors on the order of 1 pixel,
- o Sub-pixel offset estimation, followed by interpolation and image resampling, resulting in residual fractional pixel errors less than the true sub-pixel offsets.

Three main types of approximate registration methods are discussed in this section. They are area-based, contour-based, and combined methods using both areas as well as boundaries for matching.

3.1 General Issues and Strategies

We will classify feature matching techniques as region-based or contour-based, and an overview of methods is presented in sections 3.1.2 and 3.1.3 respectively. Particular new methods developed for the present effort appear in sections 3.2.1 and 3.2.2 respectively. Section 3.1.1 presents a review of the difficulties in matching dissimilar imagery.

In either case, matching requires some analytic metric for making ranked comparisons of good, poor, and ambiguous matches. The challenge is to find metrics which are robust with respect to a wide variety of illumination and viewing conditions.

One initial approximate registration method is to simply use a reference geoid and geocode both images using any navigational information, such as GPS or INS.

If map data are available, either as terrain (DTED), feature (DFAD), or both, then one approach can employ the geocoding methods described in [Kober et al,88] as an initial step for approximate dissimilar image-image registration. The rule-based "expert" assistants described in [Curlander et al,89] would then be used to select appropriate control features in both the optical and SAR images.

More accurate SAR-optical registration would then proceed on the basis of grey-level information in the two images. Of course, in the absence of any preliminary geocoding, dissimilar image registration would proceed on the basis of such pattern information alone. The following discussions concern matching based solely on pattern information within the images to be registered.

3.1.1 Overview on Difficulties in Matching Dissimilar Imagery

Precision registration of SAR and optical images is potentially difficult

because of both terrain-induced geometric distortion and radiometric differences between the two sensor types (see Fig. 1). Compounding this problem is the local decorrelation of gray values within a single SAR image. This smaller correlation length stems from the large amount of speckle appearing in SAR imagery. Such speckle effects are always present in coherent imaging systems whenever there is scene microstructure on a smaller scale than the sensor resolution.

The difficulties for registration involve both geometric and radiometric disparities. One radiometric problem source is speckle, which has been discussed in section 2.3. Two remaining major factors contributing to errors which make precise registration difficult are: nonlinearities associated with terrain and low contrast regions, local contrast reversals, and edge migration.

As discussed in [Ramapriyan et. al.,86], grey value differences among overlapping SAR images are sensitive to range variations both within and among the images, as well as incidence angle differences.

Relative geometric distortions between two overlapping images will be accentuated by rapidly varying terrain relief, resulting in compression and stretching of corresponding regions and distortions of grey level boundaries. Geometric compression actually involves the accumulation of returns from multiple terrain resolution elements into the same range cell. Because of this accumulation of returns, the radiometric value of the affected pixels can vary widely in one image from the corresponding pixels in the other image.

An extreme example of discontinuous distortion includes the phenomenon of layover. The layover effect results in certain objects which are planimetrically further distant from the sensor being imaged with a smaller range than nearby objects planimetrically closer to the sensor. This occurs because of very abrupt transitions in height, as for example with flagpoles. Such discontinuities will almost always cause errors in automated matching, unless the affected feature is very small compared to the matching window or specialized logic for detection of overlays is used.

Opposite-side effects in SAR images may be caused by changes in local incidence angles and environmental changes, both of which can result in a changed radar scattering cross-section, and result in nonlinear grey value mappings for corresponding regions. What often occurs in these cases is called "contrast reversals". These local contrast reversals are not predictable from one image to the next, and can cause problems for area correlation schemes if the affected regions are a sizeable fraction of the matching window. In fact, differences in corresponding local incidence angles can cause mismatches to occur even for manually chosen correspondences [Leberl,86], [Fullerton,86].

Low contrast regions with superimposed speckle noise and specular effects can cause problems for area-based matching metrics. In such cases, the signal-to-noise ratios (S/N) are reduced. In the case of correlation metrics, such situations often result in broad, relatively flat

correlation peaks. Therefore, the location of the correlation maximum is very sensitive in response to relatively small amounts of noise. Shadows can also be a problem for such matching algorithms, but also represent important constraints on the reconstructed terrain.

Another radiometric pathology is edge migration. This situation occurs for the direction of radar illumination just grazing the terrain, causing an abrupt transition from light to dark grey levels. As the sensor is translated in the direction of this terrain element, the position of the grazing tangent moves away from the sensor.

Matching processes depend on such grey level transitions for correspondence, either explicitly or implicitly. Edge-based matching algorithms will match such edges incorrectly. Area-based correlation methods will be affected by this effect in proportion to their sensitivity to the higher spatial frequencies corresponding to the affected grey level edge and its size relative to the size of the matching window. It is interesting to note that the direction of the edge migration effect is opposite that of the parallax direction due to relief displacement.

Even after adjusting for differing spatial resolutions, the matching of dissimilar imagery is difficult. This is because what is desired is finding the correspondences of similar object features whose image signatures are not always similar. Some of the geometric and radiometric factors which contribute to this difficulty are discussed below.

SAR and optical imagery are formed by sensing at different wavelengths. Therefore the interactions of the different wavelengths and scene roughness result in different patterns of grey level textures. This pathology is a result of the wavelength dependence of the scattering cross section coefficient σ_0 .

However, there are other dependencies of σ_0 that are important at microwave frequencies but not at optical. The discontinuity of the dielectric constant is very important for the strength of radar return from an object, but not at optical wavelengths. This is why moisture from a recent rainfall will change the SAR signature of a cornfield more than in the optical case.

Moreover, the SAR is an active, coherently imaging instrument, whereas optical sensors, with the exception of lasers, are passive. Therefore, SAR is subject to coherence-induced speckle, modeled as multiplicative noise, whereas optical images generally are subject to additive noise.

One major problem is the greater radiometric sensitivity of a SAR compared to a passive optical instrument when imaging the same scene under different viewing scenarios.

Low contrast regions with superimposed speckle noise and specular effects can cause problems for area-based matching metrics. In such cases, the signal-to-noise ratios (S/N) are reduced. In the case of correlation metrics, such situations often result in broad, relatively flat correlation peaks. Therefore, the location of the correlation maximum is

very sensitive in response to relatively small amounts of noise. Shadows can also be a problem for such matching algorithms, but also represent important constraints on the reconstructed terrain.

Certain radiometric pathologies have been observed in SAR images of lava flows, forests, and sand dunes [Blom,88]. These features have been observed to change appearance dramatically, even sometimes disappearing completely, with variations in look angle and wavelength. Subsurface imaging in some cases appears to be involved. Some statistical analyses of scatterometer data seem to indicate that discrimination between lava flows is better when look angles are less than about 35 degrees [Blom,85,86,87a]. The same type of smaller angle effect has been observed for sand dunes [Blom,81,82a,82b,87b].

The mechanisms for such behavior are not completely understood at present, and can therefore not be reliably predicted. The implication for image matching is that sometimes certain features and textures visible in an optical image of the same resolution may be missing in a SAR image of the same scene even though they are not geometrically occluded.

Another problem is that terrain-induced geometric distortions differ for both types of sensors.

Relative geometric distortions between two stereo images will be accentuated by rapidly varying terrain relief, resulting in compression and stretching of corresponding regions and distortions of grey level boundaries. Geometric compression actually involves the accumulation of returns from multiple terrain resolution elements into the same range cell. Because of this accumulation of returns, the radiometric value of the affected pixels can vary widely in one image from the corresponding pixels in the other image.

As pointed out in [Ramapriyan,86], the requirement for well-conditioned stereo intersections aggravates these problems.

An extreme example of discontinuous distortion includes the phenomenon of overlay. The layover effect results in certain objects which are planimetrically more distant from the sensor being imaged with a smaller range than nearby objects planimetrically closer to the sensor. This occurs because of very abrupt transitions in height, as for example with flagpoles. Such discontinuities will almost always cause errors in automated matching, unless the affected feature is very small compared to the matching window or specialized logic for detection of overlays is used.

3.1.2 Overview of Area-Based Matching

Template matching using any of the various forms of the correlation metric has historically been a useful method for image-image matching. This method of matching is generally region-based rather than boundary-based, although it can also be used with some degree of success for matching high-pass filtered data. Another area-based metric is the sum of absolute differences [Barnea,72].

One classification scheme of image translation estimation methods appears in [Huang,81], and consists of : Fourier-based, matching, and differential.

Differential methods apply to image pairs with small amounts of relative translation, such as successive video frames. Such methods are based on Taylor's expansion for two variables truncated to linear terms. Obviously, such an approach is not applicable to this registration problem, and will not be discussed here.

Matching-type methods refer to the use of correlation-based methods as a measure of similarity for trial areas of overlap between image pairs. The use of such methods will be discussed in sections 3.2, 3.3, 4. and 5.

Fourier-based methods refer to an explicit use of the Fourier Shift Theorem [Champeney,73], which states that if two integrable continuous functions $f(x,y)$ and $g(x,y)$ are related as: $g(x,y) = f(x + \Delta x, y + \Delta y)$ then their Fourier Transforms are related as:

$$F(u,v) = G(u,v) \exp\{-j 2\pi (u \Delta x + v \Delta y)\}$$

In practice, such methods reduce to phase correlation.

A classical method for registering images is to successively register corresponding patches within the two images using an analytic metric for comparison. For each such patch in one image, candidate trial patches in the second image are compared using this metric. The trial patch in the second image which optimizes the metric is chosen as the "matching" patch for the given patch in the first image.

Historically, the normalized correlation coefficient has been the preferred metric, although other less robust versions of correlation-type metrics have been used because of computational advantages, such as non-normalized correlation, or variations on the sum of absolute difference metrics. These three are given below:

Normalized Correlation:

$$R(m,n) = \frac{\sum_{j=1}^N \sum_{k=1}^N F(j,k) F(j-m,k-n)}{\left[\sum_{j=1}^N \sum_{k=1}^N F(j,k) \right] \left[\sum_{j=1}^N \sum_{k=1}^N F(j-m,k-n) \right]}$$

Here $F(j-m,k-n)$ refers to an offset of (m,n) with respect to $F(j,k)$.

Correlation:

$$R(m,n) = \sum_{j=1}^N \sum_{k=1}^N F(j,k) F(j-m,k-n)$$

Sum of Absolute Differences:

$$R(m,n) = \sum_{j=1}^N \sum_{k=1}^N |F(j,k) - F(j-m,k-n)|$$

Normalized correlation is the preferred method whenever computational cost is not an issue. The value of the normalized correlation coefficient is bounded absolutely, and always lies between -1 and +1. Therefore, absolute values of this metric close to +1 achieved for trial offsets are close to local maxima in the correlation surface. The normalized metric can allow for a constant multiplier between the gray-values for corresponding pixels in the two images. Also, the normalized correlation coefficient theoretically achieves its maximum value, in the absence of noise and nonlinearities, at the correct offset between the two images.

This is not always the case for non-normalized correlation algorithms, which can achieve high correlation values at some incorrect offsets simply because of high gray-values at certain localities. The sum of absolute values of differences generally performs better, in the sense of acceptable accuracy, than the non-normalized correlation approach, but also not as well as normalized correlation [Svedlow,77].

Correlation is computationally expensive. One route toward reducing the amount of computation has been to perform the equivalent operation in the Fourier domain, using the Fourier Convolution Theorem [Champeney,73]. The basis of computational savings using the FFT is that the latter requires a computation asymptotically proportional to $N \log N$ rather than N^2 , where N is the dimension of a square image. For relatively large images, there can be considerable savings using the FFT approach. Up to an order of magnitude or more efficiency can sometimes be achieved, depending on the relative sizes of the convolution block L and image size N [Pratt,78].

However, this approach does not deal with the observation that the determination whether a certain offset is very incorrect should somehow require less computation than the determination of a correct offset. This view is at the root of the idea of recursively searching for the correct offset in a "pyramid" of increasing-resolution versions of the image pairs [Rosenfeld,84].

This point-of-view is also the basis of the sequential approach using the sum of absolute differences metric in [Barnea,72]. Here, for any trial offset, whenever a pre-determined threshold has been exceeded before the entire sum has been evaluated, the summation is suspended and a new trial offset is evaluated.

Hybrid algorithms using this approach for a rough estimate of offset, followed by a version of normalized correlation using statistical pre-processing have been suggested [Pratt,73]. Clearly, this idea can be generalized to the use of other robust metrics following the initial rough estimate.

Another problem with correlation measures in general, including normalized correlation, includes the broad, flat nature of the peak regions in the correlation surface. This characteristic has advantages and disadvantages. An advantage of broad peaks is their "pull-in" range for search techniques that don't sample everywhere or that employ averaging, as in multi-resolution processing [Rosenfeld,84].

A disadvantage is that registration accuracy can be compromised whenever peaks are not sharp, and smaller perturbations due to noise can potentially have greater effects on accuracy.

This broad peak characteristic of correlation techniques occurs because spatial relationships in images are ignored for the most part. Instead, the values of the correlation metric are really related to energies of the broad areas within the images, i.e. the energy content of the low-frequency portions of the images. Since phase shifts of these lower frequencies can be relatively large compared to the pixel resolution and still be relatively small compared to the corresponding low-frequency wavelengths, there is a lack of sensitivity in correlation metrics to phase shifts. Therefore, correlation surfaces tend to have broad peaks.

One approach toward solving this problem has involved the preferential use of phase information in the images. The idea here is that the Fourier phase content of the images contains more accurate information than the low-frequencies which dominate correlation metrics. This is the basis for phase correlation methods [Kuglin,75,79], [Pearson,77], [DeCastro,87].

Using the previous example of the functions $f(x,y)$ and $g(x,y)$ related by translations Δx and Δy , the inverse Fourier transform of $\{F(u,v) / G(u,v)\}$ is just the inverse transform of the phase term, and thus is equal to the Dirac Delta Distribution evaluated at Δx and Δy , i.e. $\delta(\Delta x, \Delta y)$

Now, because of the effects of sampling and finite image sizes, sidelobes occur in addition to a main peak. Therefore, in practice, this technique reduces to correlation in the Fourier phase domain.

Such methods are also capable of subpixel accuracy with the use of interpolation, but suffer from the problem of potentially high sidelobes. Such a correlation surface, with sharp peaks and high sidelobes, does not lend itself easily to hierarchical processing with reduced resolutions because of the narrow "pull-in" range of the main lobe. However, used in conjunction with other methods which can acquire the main lobe, phase correlation can be a useful method for refining initial offset estimates to subpixel accuracy.

Another approach to rectify this broad peak problem has been to concentrate on the edge content in the images. This approach has both intuitive appeal as well as some theoretical justification.

Intuitively, it would seem that the "meaningful" information in an image lies at the locations of large, structurally significant contours and boundaries. Particularly in SAR images, the smaller edges are more often

due to noise, speckle, and imaging effects. However, the larger edges are due to terrain and thematic value changes. It would seem that it is these edges and boundaries that should be involved in a registration scheme.

This notion has been explored on a more analytical basis from the standpoint of "optimally" filtering the images as a pre-processing step prior to registration. One approach has been to decorrelate the images by applying a "whitening" filter [Pratt,73], [Svedlow,78]. These methods essentially differ in their assumptions concerning image structure and statistical properties. The conclusions of these works point toward the use of pre-processing filters which can be approximated, under certain assumptions, by gradient filters [Svedlow,78], or Laplacian filters [Pratt,73].

Of course, there are some problems associated with this approach also. The high-frequency edge content of an image is relatively small compared to the total area of an image. Edges can be broken up slightly differently in two images which otherwise contain no other perturbations. Edges may also have slight variations in thickness. Therefore, such perturbations can lead to misalignment sensitivities for algorithms, like edge correlation, which examine the degree of "match" in overlapping edge images. One approach is to condition the edges in both images. This would include normalizing their intensities and broadening them. However, broadening edges can lead to reduction of registration accuracy.

These considerations mentioned above suggest that correlation-based matching involves considerable searching and is computationally expensive, unless the hierarchical "coarse-to-fine" approach previously discussed is used. The capabilities of correlation algorithms for achieving high accuracies often seems to require some form of pre-filtering to enhance edge content. However, certain potential instabilities are involved with the use of edge correlation.

The use of correlation, however, assumes that there is little nonlinear geometric or radiometric distortion between the template and the actual imaged feature instance [Lahart,70].

3.1.3 Overview of Boundary-Based Matching Methods

One method of boundary matching in binarized images which has achieved some success is "chamfer matching" [Barrows, 78]. This method uses a distance array between features in patches being matched, and estimates the translations between patches by searching for those offsets which minimize the distance array sums. This technique generally requires cleaner extraction of edges than does edge correlation. As is the case with all edge-based matching techniques, this metric is more sensitive to perturbations perpendicular to lines in the image than parallel to such lines.

One method developed at VEXCEL [McConnell,87] performs matching for translation estimation by creating binarized images from Marr-Hildreth operator zero crossings. This method has the advantage of accurate computation of edge content using this second-order operator, as well as

the stability of regional matching. This method is especially well-suited for registering opposite-side satellite SAR images.

Another method was developed at VEXCEL for registering ice floes in arctic SAR imagery. This method performed boundary matching using dynamic programming, and reduced a two-dimensional search problem to one involving one-dimension. This method employed the psi-s representation of closed contours [Ballard,82], which allowed a convenient representation for translation and rotation. The particular implementation of the dynamic programming algorithm [Sankoff,83] was that used for string search and other sequence comparisons. The assumption was that the ice floes are rigidly rotated and translated, but some very localized distortions such as expansion, contraction, insertion and deletion could be tolerated.

Graph matching is a method which incorporates topological relationships into the matching process without undue emphasis on "exact" metric correspondence. This method has found considerable use in image-image matching, for example [Price, 82], [Ballard,82], [Nevatia, 82], [Shapiro,81], [Ayache,87]. Such problems involve subgraph isomorphism and can potentially be complex. Although graph matching problems belong to the worst-case computationally intractable class NP-complete [Aho,74], the use of heuristics to reduce the search has been successful.

In fact, all computationally successful graph matching algorithms must somehow use some means of distinguishing salient features in order to avoid the problem of combinatorial explosion when searching.

For example, in [Medioni,84], graph matching was used for both image-image and image-map matching. The computational complexity of matching was reduced by the use of a "coarse to fine" matching strategy which extracted isolated, long edges first, and matched these to the model at low resolution using relaxation.

Other boundary-oriented graph matching techniques have emerged from the field of robot vision. Typical of the problems encountered in this field is to correctly identify the individual parts in an imaged pile of parts which may be overlapping and are partially obscured by each other. Since the surfaces of these parts are often smooth, the boundaries contain most of the identifying information.

An example of such work is [Bolles,82]. Significant local features are identified, such as corners and holes. These features are then clustered, and matching proceeds "cluster to cluster". The algorithm attempts to find the most significant cluster. A graph matching formulation is created wherein the nodes are matches of image and model features, and edges are pair-wise assignments between nodes. The match problem is then equivalent to searching for maximal cliques, which has complexity NP-complete. This method assumes that features will be clustered closely together.

Another method is that found in [Ayache,84]. Here, polygonal approximations of extracted boundaries and model sides are matched using a strategy of matching the longest sides first, as is done in [Medioni,84]. Lengths of sides and corner angles are used in measures of

"compatibility". This method also attempts to generate "hypotheses" and continues until a sufficient number of hypotheses are evaluated and a good match score has been obtained.

The method of [Turney,85] divides a template of length n into $n/2$ subtemplates of length h . Every other pixel on the boundary of the template starts a subtemplate. Each subtemplate of each model object is then compared to the extracted object boundary in the sensed image.

The boundaries are represented using a variation of the psi-s representation mentioned previously. The matching metric is the least-squares measure of the differences. The method attempts to reduce the combinatorics of the comparisons by selecting "most salient" subtemplates. The successful matching of such salient subtemplates is weighted more than other matches. Again, this algorithm is slow.

The spirit of this last method, however, is carried out with significant computational improvements in [Schwarz,85]. Moreover, [Kalvin,86] continues computational improvements on this approach by employing the clever idea of "geometric hashing". This hashing is accomplished by the use of a "footprint" of an object's boundary. Only a small number of object footprints are retrieved for potential matching checks. This number of retrieved footprints does not depend strongly on the number of objects in the model database.

A footprint is a crude geometric characterization of an object boundary using 5 dimensions. A footprint is generated by a mapping that is not necessarily 1-1, but satisfies invariance under translation and rotation. continuity is essentially preserved by having locally similar objects into similar footprints. The five dimensional representation of each point of the object's boundary involves the first four Fourier coefficients of the boundary and the turning angle of a polygonal approximation at that boundary point.

The hashing occurs because 5-D space is divided into hypercubes of fixed size, and for each hypercube there is a list of all models whose footprints lie in that hypercube. In this way, searching for potential matches of boundary segments is considerably reduced.

Another approach is the method of invariant central moments. The use of this method as a theoretical pattern recognition invariant under translation, scaling, and rotation first appears in [Hu,62], and was used in [Wong,80] for SAR-optical matching. The latter approach essentially used orthogonal expansions via the first seven invariant moments of a radar sub-image and a trial offset window in the optical window as the components of feature vectors to be correlated:

$$R(x,y) = \{ \sum_{i=1}^7 M_i N_i(x,y) \} / \{ \sum_{i=1}^7 M_i^2 \sum_{i=1}^7 N_i^2(x,y) \}$$

where: $M_i = i^{th}$ invariant moment of the target SAR sub-image

$N_i(x,y) = i^{th}$ invariant moment of the optical window at trial offset (x,y)

Now, ordinary moments are defined as:

$$m_{pq} = \int_{-\infty}^{\infty} f(x,y) x^p y^q dx dy$$

where: $f(x,y)$ = image grey value at (x,y)

These ordinary moments were used in [Hu,62] to define central moments, which were then used to define the seven invariant moment functions used above, M_i : $i=1, \dots, 7$. A disadvantage of using the higher invariant moments is their instability with respect to noise and discretization effects, as well as integration errors. Moreover, moments are not contrast invariant [Maitra,79].

Hierarchical matching using a resolution pyramid can decrease the amount of computation considerably if the important features do not disappear when the resolution is decreased.

3.2 Methods Used in Present Effort

Three types of approximate registration methods are discussed in this section: area-based, contour-based, and combined. These are discussed in the following 3 sub-sections.

3.2.1 Area-Based

Two algorithms were formulated and tested for estimating pixel offsets of an optical chip within a SAR image. Both of these methods assumed only translational differences between a small optical chip and the corresponding region in the SAR image. No appreciable terrain distortions were present.

One method involved the Karhunen-Loeve Transform (K-L), and the other concerned the use of the Maximum Noise Transformation (MNF). These two methods are detailed in the next two sections, 3.2.1.1 and 3.2.1.2.

3.2.1.1 Translation Estimation Using the Karhunen-Loeve Transform

The following intensity transformation for SAR images for use in matching with an optical chip is based on the Karhunen-Loeve (K-L) Transformation. This orthogonal transformation has been used previously as a pre-processing step for both noise removal, and for intensity re-mapping a SAR image on the basis of intensities in an optical image with which the SAR image has been registered previously.

However, the K-L transformation was used in the following way for registering a SAR image and an optical chip:

A chip in the SAR image was re-mapped using two distinct intensity re-mapping functions, one of which was registration independent while the other was registration dependent. The registration independent method used involved histogram equalization [Castleman,79]. The registration dependent method used the eigenvector of the maximum eigenvalue of the K-L

transformation of the covariance matrix of the pixel co-occurrence values for a given trial registration.

Therefore, for each trial offset of the optical chip within the SAR image, the two resulting intensity re-mapped versions of the corresponding SAR chip were evaluated using normalized correlation. The idea was that at the correct offset, there should be a greater similarity between these versions than at other offsets.

This hypothesis proved to be generally correct only in a local sense, ie. the correct offset was usually found to be associated with a local maximum of the resulting correlation surface. However, it was observed that at those "correct" local maxima which were not also global maxima, there was a much more dramatic local change in the mean intensity levels of the K-L re-mapped SAR image. These results are discussed in [Curlander and Kober,89]. Therefore, the combination of these two criteria is a promising method for further investigation.

This algorithm is briefly described below:

For a trial registration offset of the optical chip in the SAR image, compute:

- o Form the (2x2) composite covariance matrix K corresponding to a sum of the covariance matrices for each pixel pairing in the window. Here K is given by:

$$K = (1/N) \sum_{i=1}^N (\underline{f}_i - \underline{m})(\underline{f}_i - \underline{m})^t$$

$$\text{where } \underline{m} = (1/N) \sum_{i=1}^N \underline{f}_i$$

$$\underline{f}_i = (f_1^{(i)}, f_2^{(i)})^t \quad i=1, \dots, N$$

$i = i^{\text{th}}$ registered pixel, $i = 1, \dots, N$

$f_1^{(i)}$ = grey value of i^{th} registered pixel in image 1

$f_2^{(i)}$ = grey value of i^{th} registered pixel in image 2

- o Compute the maximum eigenvalue λ of K and its eigenvector $(x_1, x_2)^t$:

$$\lambda = \{a+d+[(a-d)^2+4b^2]^{1/2}\}/2$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b/[b^2+(\lambda-a)^2]^{1/2} \\ (\lambda-a)/[b^2+(\lambda-a)^2]^{1/2} \end{bmatrix}$$

where the coefficients a,b,d are entries of the matrix K:

$$K = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

- o For each SAR pixel in the trial offset window, compute a new grey value g_i :

$$g_i = x_1 f_1^{(i)} + x_2 f_2^{(i)}$$

- o Compute the normalized cross correlation of the intensity remapped SAR window with the same patch in the SAR image which has been intensity remapped using histogram equalization.

3.2.1.2 Translation Estimation Using the Maximum Noise Fraction Transformation

Another intensity transformation in the spirit of orthogonal decomposition is the Maximum Noise Fraction Transformation (MNF) [Green et al,88]. This transformation is a generalization of the K-L (or principal components) transformation for multi-spectral imagery.

The conventional wisdom regarding the K-L transformation of a multi-spectral data set is that it can be used to provide a sequence of "eigenimages" of decreasing S/N, and that the first two or three incorporate most of the important information in the dataset. In fact, it is shown in [Green et al,88] that this is true only if the noise components in all of the bands are uncorrelated and of equal variance. An example is given whereby the K-L transformation specifically does not provide an optimal ordering for image quality in the case airborne thematic mapper (ATM) simulator data.

However, the MNF transformation is specifically derived to achieve increasing image quality, as measured by S/N, for decreasing component number. The MNF transformation can be used to order a new sequence of images into increasing S/N quality, so that the lowest quality images can be subjected to the most intense averaging procedures of noise removal. The MNF transformation is then inverted to return to the original multi-band data set, but which has now been noise filtered.

In the special case wherein all of the noise resides in one band only, there is a special solution to this procedure. This is approximately true in the case of an optical chip and a SAR image, since the SAR image suffers from a great deal of speckle.

Therefore, in the spirit of the previous algorithm in section 3.3.3.1.1, this transformation was applied to the SAR image for each trial offset of the optical chip, followed by normalized correlation. The results, discussed in [Curlander and Kober,89], were slightly better than for the K-L re-mapping method.

This algorithm is briefly described below:

- o Set $z_1(x) = \log_{10} R(x)$

where $R(x)$ = grey value in SAR image at pixel x

- o Set $z_2(x)$ = grey value in optical image at pixel x
- o For each offset of optical window in SAR image, form:
 - m_1 = mean value of $z_1(x)$ in window
 - m_2 = mean value of $z_2(x)$ in window
 - $\Sigma = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$ the covariance matrix of the pixel co-occurrences
- o Set $z^*(x) = m_1 - (v_{12}/v_{11})(z_2(x) - m_2)$
- o Form new re-mapped SAR window consisting of 10^y for each pixel, where
 $y = z^*(x)$
- o As an alternative to the exponentiation step above, the use of histogram hyperbolization [Frei,77] was used. This procedure is briefly described below.

Histogram hyperbolization represents a modification to the idea of histogram equalization for contrast enhancement which takes the human visual system into account. The idea is to create a uniform distribution of "perceived" intensity values, as distinct from a uniform distribution of actual values (histogram equalization). Empirically, the use of this procedure on logarithmically processed imagery has resulted in better contrast enhancement than with re-exponentiation.

A logarithmic response to image intensities, corresponding to perceived intensities, can be approximately modeled as:

$$B = \log(J(I)+c)$$

where: I = intensity value prior to intensity re-mapping

$J(I)$ = intensity value after intensity re-mapping

B = perceived brightness value

Equalization of perceived brightness values leads to the expression:

$$J(I) = c(\exp [\log(1+1/c)F_I] - 1)$$

where: F_I = cumulative distribution function of the image intensities prior to re-mapping.

c = arbitrary constant

Another direction for SAR-optical image matching involves concentrating on the higher frequency portions of the datasets. Intuitively, the edge

contents of both images have more in common than do the lower frequency, broader regions. As an alternative to edge enhancement methods, the following (3x3) spatial de-correlation filter [Pratt,78] was used:

$$\begin{bmatrix} \rho^2 & -\rho(1+\rho^2) & \rho^2 \\ -\rho(1+\rho^2) & (1+\rho^2)^2 & -\rho(1+\rho^2) \\ \rho^2 & -\rho(1+\rho^2) & \rho^2 \end{bmatrix}$$

This filter was used both by itself, and with both the K-L and the MNF re-mapped versions of a SAR chip prior to normalized correlation. Results are discussed in [Curlander and Kober,89].

3.2.2 Contour-Based

Any algorithm that segments data from an EO image to cue a search in a SAR image must take into account the differing projections for both images. Projection effects, due to the angle between the sensor look direction and the feature as well as the sensor depression angle, are important determinants in the dissimilarity between an corresponding SAR and EO image pair. Such projection effects are less of a problem for matching long linear features than shorter areal features.

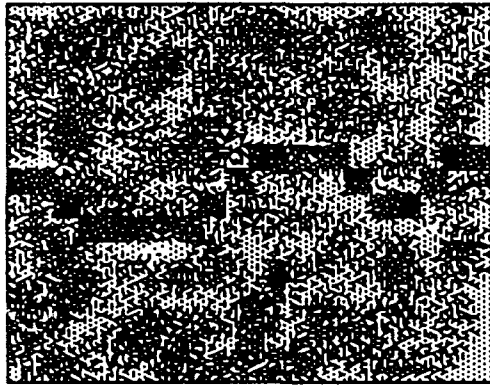
3.2.2.1 The Profile Method For Contour-Based Matching

One new algorithm developed at VEXCEL searches for long, linear features characterized by local contrast on both sides of the feature. The targeted features are only a few pixels wide and are locally brighter or darker than a surrounding area of size defined by the user. Essentially, the algorithm uses a summed profile of the pixel values in a local region as shown in Fig. 3-1.

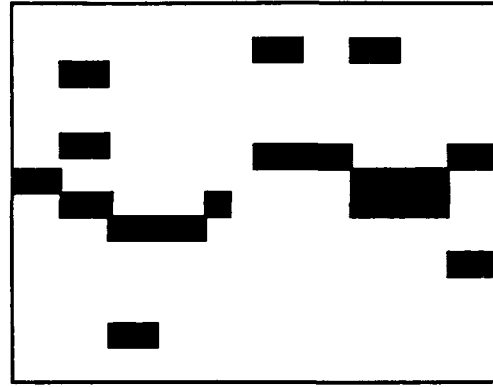
In a small window, rows and/or columns of pixel values are summed to produce a profile. Ideally, any horizontal or vertical linear feature that is only a few pixels wide compared to the band size of the profile will stand out by having a sum value significantly greater than, or less than, the mean value of the profile. This idea exploits the local contrast of the edge.

This method is also more sensitive than many other methods of edge segmentation for regions with narrow edges, ie. having a thickness of one or two pixels, superimposed on a broad, uniform background. The method's robustness to noise effects is improved because integration, in the form of adding pixel values, is a more stable operation than the usual differentiation step in edge segmentation operations. Searching preferentially for longer horizontal lines than vertical lines can be accomplished by using a local region mask which is more elongated in the horizontal direction.

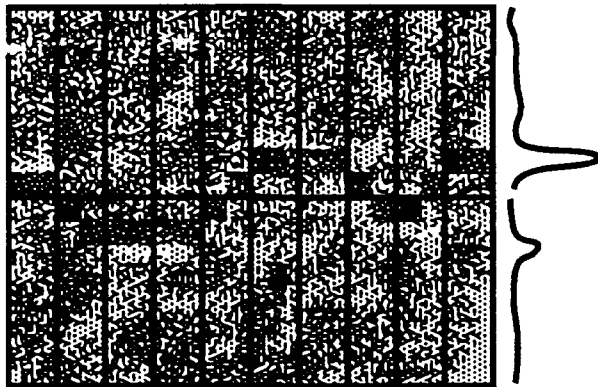
Repeatedly applying this method to multiple, adjacent windows in an image results in more complex segmentations. Rows and columns which were flagged as outliers in their local profiles are compared with adjacent rows and columns in the neighboring windows. If the adjacent windows contain flagged pixels in the adjacent rows and columns, those pixels are kept. Otherwise, the flagged pixels in the original window are discarded.



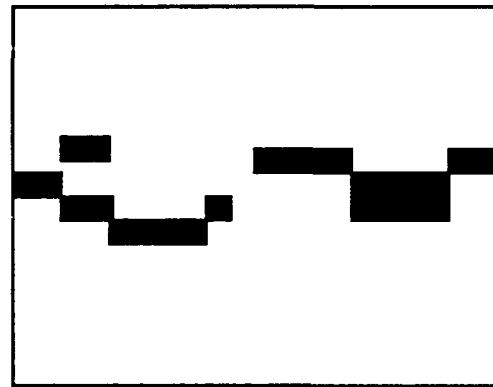
(a)



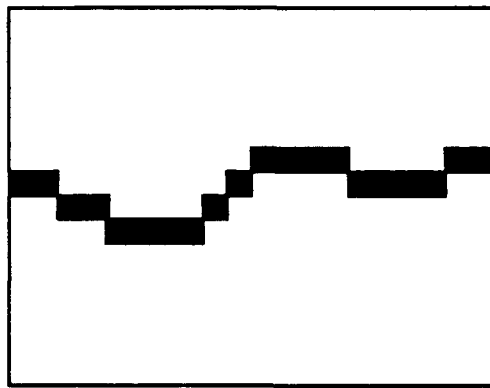
(c)



(b)



(d)



(e)

Fig. 3-1: (a) A feature of interest in a noisy background, (b) window divided for local summed profiles in vertical direction, (c) edges extracted by thresholding profiles in each block using local statistics, (d) edges after removal of discontinuous segments, (e) vector representation of edge contour.

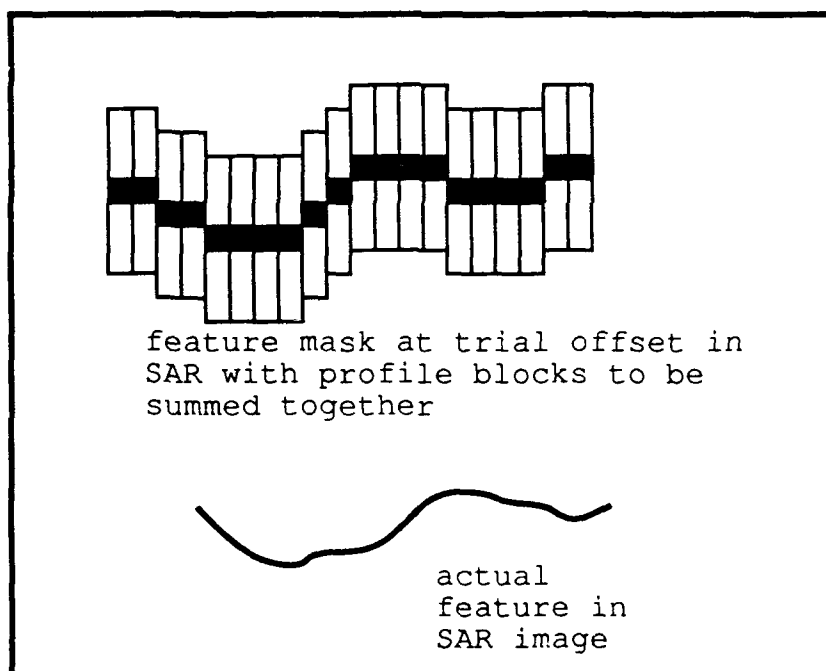


Fig. 3-2: Use of edge contour from optical image to cue customized profiling in SAR image. (Note when the slab produced by the profile is laid around the vector contour is exactly placed upon the real contour in the SAR image, the profile will display a peak at the center).

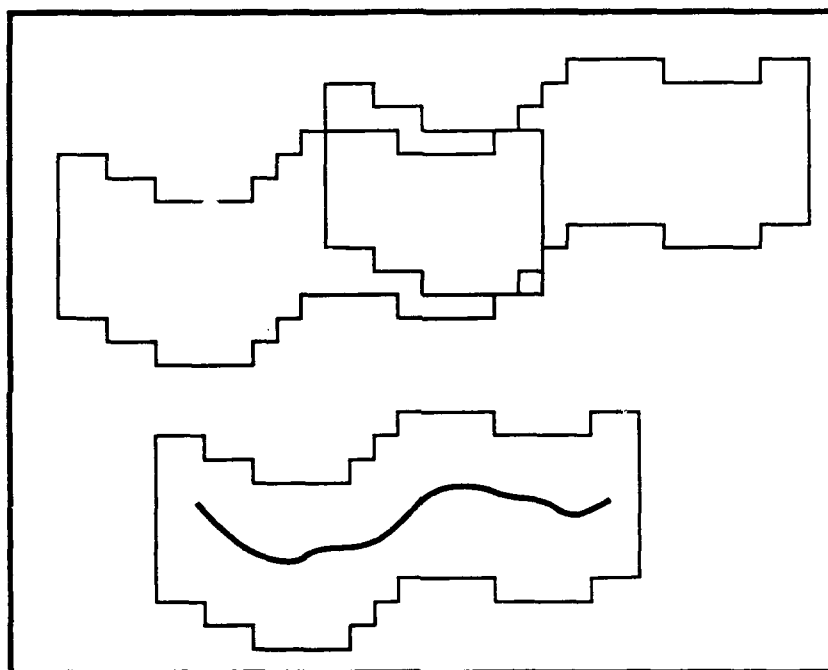


Fig. 3-3: Slab to be sum-profiled at various offsets in the SAR image. At each offset any outliers in the profile are highlighted and added to the output image at the corresponding locations indicated by the peak in the profile. The actual locations of the contour in the SAR may be detected by several offsets of the slab, but will always highlight the same locations in the output image. The vectorized mask can then be correlated with the output image to indicate the best offset in the SAR.

In this algorithm, it is important to ensure that whenever a potential edge is discarded, it cannot cause another edge in an adjacent edge to also be discarded and thereby cause a cascading sequence of edge discards.

A subsequent process can gather the flagged segments together based on a closeness threshold for the adjacent endpoints. The output of this process would be a series of linked lists, each consisting of a "streak" of some length. The choice of which streaks to use for matching can be based on streak length.

A similar algorithm is applied to the SAR imagery, with suitable modifications taking into account the characteristics of SAR. Once a rough estimate of the location of particular feature in the SAR is known, then a variety of methods may be used to refine this initial estimate.

One variation of the above profile method is now described (see Fig. 3-2). This algorithm uses the same summed profile, but only for streaks generated from the EO imagery. The corresponding streak in the SAR can lie anywhere within a region of uncertainty. An uncertainty in rotation is also present. An assumption of $\pm 3^\circ$ is reasonably consistent with present INS drift rates and mission durations of a few hours.

At a candidate offset location within the SAR image, cued by some streak, a "total profile" is formed by summing successive profiles centered at pixels along the streak with respect to one of its endpoints. Next, the total profile is thresholded at N standard deviations above and below the mean.

Any pixels in the total profile which are identified as outliers indicate streaks parallel to the currently sought streak in the EO image at an offset corresponding to the location within the profile.

This procedure is repeated for trial offsets in the SAR image within the region determined by the uncertainties in range and azimuth (see Fig. 3-3). The number of standard deviations associated with each parallel streak is summed into the output image to produce a surface. Following these computations for all trial offsets, the target streak from the optical image is correlated with this derived surface. Because of the binary nature of these streaks, a sum of absolute differences metric can be used as the correlation metric.

3.2.2.2 "Beam Thinning"- A Future Contour-Based Matching Algorithm

Another matching algorithm was also developed theoretically, and is intended to be investigated during Phase II for matching intensity contours which have been distorted by terrain. Of course, such terrain-induced distortions will be different for SAR and optical sensors. Therefore, any successful matching technique must be robust with respect to metric distortions that represent differences between sensors, while still retaining sensitivity to similarities arising from the common terrain imaged.

The algorithm described briefly below represents an algorithm which essentially can represent the local properties of a curve which are more "topological", as well as metric descriptions. In particular, it can adequately represent the regions of changing curvature at various scales of resolution.

It is believed that curve descriptions which contain qualitative information such as the sequences of positive and negative curvature, aside from their magnitude, will be important for matching images with sensor-dependent distortion. This is because the sensor type will generally modify the magnitudes of such local curvatures, but will not change their sign patterns.

This curve representation method, called "beam thinning", works using a hierarchical description of a curve in which increasing detail appears with increasing depth of the describing graph structure. The hierarchical description for extracted curves is the following:

Approximate the curve by a "slab" of thickness d_k as shown in Fig. 3-7. For each thickness d_k , the curve is easily partitioned into convex and concave regions. For each such concave/convex portion, there is a node in the descriptor graph. Such a node contains descriptive information for that portion of the curve, such as the curvature and the length.

As the thickness of the approximating slab decreases, each convex/concave region can potentially split into concave/convex subregions. This can be expressed in graph theoretic form by forming subnodes under each of the nodes. In this way, there is a description of the curve for a sequence of slab thicknesses $\{d_n\}$ by means of a graph whose node levels correspond to the $\{d_n\}$. The ultimate level of descriptive detail corresponds to a slab thickness of 1 pixel (see Fig. 3-4).

Such a description allows a descending sequence of approximate curve matchings to occur based on a succession of increasingly detailed trends in the curve description. This is one way to deal with the problem of insertions and deletions of pixels, as well as small-scale stretching and dilation of a curve.

Since for each approximation level d_n there is a corresponding level in the descriptor graph, dynamic programming can be used for matching convex/concave regions just as it can on the individual pixel level.

An issue is the precise computation of an approximating slab of thickness d for a curve C . An appealing heuristic is the minimization of the total mechanical strain energy in the slab:

$$\int_0^L \kappa^2(s) ds \quad \text{where: } L = \text{length of the curve } C$$

ds = element of arc length

$\kappa(s)$ = curvature

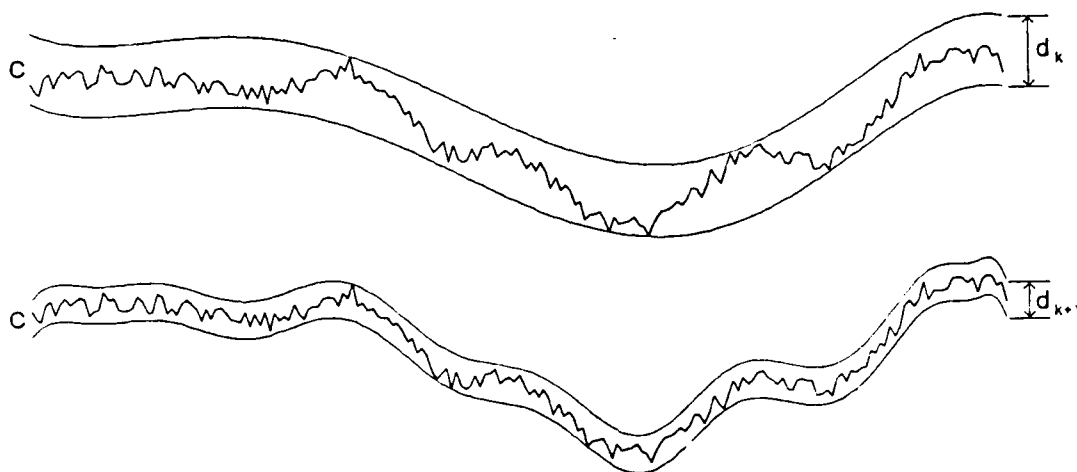


Fig. 3-4: "Beam thinning" decomposition of an image curve

However, the previous measure is for curves, not beams. Moreover, it is scale-variant.

A better measure of stress energy for curves is the scale-invariant measure proposed by [Bruckstein and Netravali,90]:

$$C^* \{ \Psi(s) \} = L \int_0^L k^2(s) ds = L \int_0^L [d\Psi(s)/ds]^2 ds$$

where $\Psi(s)$ = the Ψ -s representation of the curve C, ie. the tangent angle as a function of arc-length s.

This measure will have to be extended to a beam of non-zero thickness d.

Another algorithmic problem concerns the determination of the beam such that it covers the given curve subject to minimizing this scale-invariant energy measure.

Finally, there is the possibility of using eigenvector decomposition for further hashing of the "feature vector" corresponding to the curve description at any level of the descriptive hierarchy graph.

The proposed method has the advantage of providing a "coarse-fine" representation that is gaining acceptance for image analysis in general. However, the choice of a robust computational method to achieve this has been a problem in the past. Fourier-based and other globally-symmetric,

orthonormal basis function methods tend not to be responsive enough locally in a compact form.

The proposed method is also stable enough to allow the use of convexity/concavity concepts which, in the past, have been unstable descriptors for noisy curves.

3.2.3 Combined

An appealing approach is to combine the area and contour-based methods described in the previous sections into a procedure which exploits the strengths of each.

It has been found that the use of correlation methods described in section 3.3.1 depends heavily on knowledge that the images to be mapped are rotationally aligned. The remapping and correlation algorithms are very sensitive to residual rotational errors between the images.

Vector-based methods, however, can be used to indicate and quantify the amount of residual rotation present between the SAR and the optical images. Thus the vector-based contour methods could be used initially on an EO-SAR image pair to determine the amount of rotation between the two images. This method could be used at several locations in any given image. From the rotational offsets and the angle of the vectorized contour at each location, an origin and angle of rotation can be estimated for the entire image set. One of the images can then be rotated and resampled to match the other. The correlation-based matching metrics can then be expected to provide improved performance.

3.3 Results

The results of the SAR-optical image registration experiments are discussed in the following subsections. The results of area-based methods are described in section 3.3.1, while contour-based streak detection and matching methods appear in section 3.3.2.

3.3.1 Area-Based Matching Results

As previously described in [Curlander and Kober, 89], a full range of tests comparing the Maximum Noise Fraction Transformation and Karhunen-Loeve remapping methods with more conventional algorithms as normalized correlation and edge mapping for SAR-optical registration. Table 3-1 shows the results of these tests on 15 image sites in four Landsat (Band 4)-SEASAT image pairs. Each region in the SAR image uses 80x80 pixels and each region in the IR images uses 60x60 pixels. In addition to the Yuma data set shown in Fig. 2-1 and 2-2, similar data was available for the Altamaha River GA region and the Wind River Basin WY region [Curlander and Kober, 89].

SAR to Optical Image Matching Tests

Test Area	SAR Source	Optical target	n	h	Correlation Coefficients at expected peak							Correlation Coefficients at actual peak									
					k	e	l	m	p	y	o	n	h	k	e	l	m	p	y	o	
1	Yuma	137, 370	147, 380	9.5e-4	.0871	.0576	.1573	.0382	.0208	.8131	.0841	.1465	141.370	137.370	137.390	140.370	137.373	137.371	152.370	131.373	139.371
2	Yuma	371, 227	381, 237	.0940	.0837	.9591	.0215	.1067	.1067	.3594	.0250	.1242	189.238	190.238	190.236	186.229	185.239	182.238	176.234	189.239	190.243
3	Yuma	1, 171	1, 181		.0526	.4287	.0531	.9812	.8847	.9289	.8450	.0436	.1109	.1011	.9651	.1202	.1448	.1530	.0649	.1898	.2408
4	Yuma	390, 84	400, 94	.1035	.0531	.9369	.0339	.1367	.0973	.0913	.0691	.0026	20.179	20.179	15.180	12.171	10.188	21.185	1.171	1.175	2.172
5	Yuma	360, 420	370, 430	.0322	.0280	.8714	.0109	.0373	.0385	.4407	.0598	.1681	.0992	.0685	.3313	.0833	.9996	.9193	.9729	.8489	.2953
6	Yuma	275, 328	285, 338	.2066	.1977	.5893	.1875	.2698	.2410	.0343	.2377	.0453	390.84	390.84	410.84	390.98	392.95	393.95	408.84	393.87	409.85
7	Alta	269, 128	279, 138	.4155	.4484	.9354	.0252	.7862	.7108	.3572	.6882	.3409	.1582	.1186	.9415	.1020	.2208	.1960	.4068	.2191	.1691
8	Alta	145, 162	155, 172	.3459	.4321	.9561	.0691	.8906	.8183	.1038	.8805	.2743	177.420	163.424	162.401	160.435	177.425	177.425	175.433	177.425	179.429
9	Alta	91, 348	101, 358	.2347	.3371	.7188	.0174	.6708	.4087	.0939	.3177	.1990	.0916	.0813	.8775	.0776	.1352	.1277	.7709	.1404	.3728
10	Maha	35, 27	45, 37	.3881	.3805	.9538	.2996	.9276	.9682	.3894	.8519	.018	290.334	290.328	275.328	275.349	290.333	290.333	276.328	290.333	276.328
11	Maha	104, 180	114, 190	.4757	.4595	.9662	.3105	.9235	.9185	.3869	.8796	.4851	.5676	.6145	.9478	.1202	.9300	.8490	.7239	.8317	.6190
12	Maha	165, 260	175, 270	.2728	.2759	.8147	.4111	.9859	.9609	.2538	.9253	.3036	.145.174	.145.180	.158.173	.166.178	.153.162	.156.181	.157.173	.153.162	.158.174
13	Wind	17, 354	27, 364	.0491	.0442	.8608	.2023	1.000	.9596	.0230	.8725	.0928	.4279	.5103	.9604	.2482	.9998	.9746	.8445	.9334	.4488
14	Wind	132, 241	142, 251	.6299	.7310	.9595	.5638	.9932	.9481	.4896	.9460	.5734	.112.353	.112.358	.104.348	.104.348	.104.348	.104.348	.104.348	.99.359	.104.349
15	Wind	238, 97	248, 107	.0256	.0115	.9334	.2168	.0184	6.6e-4	.1777	.6471	.0476	.2553	.3506	.7341	.0761	.6923	.4744	.5112	.4087	.3120
													56.48	56.48	36.28	55.27	46.47	46.47	41.36	42.47	42.37
													.4396	.4240	.9599	.4066	1.000	.9612	.5847	.9298	.5180
													125.198	125.190	123.200	105.181	105.200	104.199	108.194	122.181	109.191
													.4931	.4699	.9662	.3317	1.000	.9744	.7297	.9361	.6053
													186.281	186.281	174.260	165.267	166.260	169.260	168.271	168.261	169.272
													.3444	.3461	.9469	.4514	1.000	.9738	.6062	.9312	.4361
													17.354	17.354	17.374	38.354	38.367	36.354	17.358	36.356	19.358
													.2783	.1801	.8845	.2467	1.000	.9698	.2567	.8998	.1714
													153.263	153.262	142.250	143.269	132.261	147.248	143.512	132.261	144.252
													.7754	.9905	.9597	.6549	.9982	.9624	.6527	.9512	.6010
													235.106	247.97	253.111	252.98	239.116	219.101	252.98	237.99	257.99
													.0641	.0671	.9433	.7818	.1492	.1372	.5703	.1486	.2626

n = normalized correlation
 h = histogram matched SAR image to optical image - normalized correlation
 k = k-l remapped SAR image to histogram remapped SAR image - norm corr.
 e = normalized correlation of reduced resolution edge images of SAR and optical
 l = normalized correlation between k-l remapped log SAR and optical imagery
 m = normalized correlation of m-l remapped SAR and optical
 p = normalized correlation of the decorrelated m-l remapped SAR and the decorrelated optical
 y = normalized correlation of m-l remapped SAR and histogram hyperbolized optical imagery
 o = normalized correlation of the decorrelated SAR and the decorrelated optical

Table 3-1:

The Karhunen-Loeve (K-L) method utilizes a normalized correlation of the K-L remapped regions of the SAR image with histogram remapped regions of the optical image. The histogram remap is a global operation, while the K-L transformation is a local operation. The normalized correlation is then operating on two images, one of which changes with each candidate offset while the other remains constant.

It has been found that the correlation between the K-L remapping and the histogram remapping produces sharper, more distinct correlation peaks than are produced by normalized correlation or usual methods of edge matching.

An example of these results is shown in Fig. 3-5 for a region of the Altamaha River GA region, for which there was also collateral coverage from Landsat TM and SEASAT SAR.

The Maximum Noise Fraction Transformation (MNF) method uses a normalized correlation of the MNF transformed SAR image and the optical image. For a given optical image chip, the MNF transform of a corresponding SAR image chip is performed for each trial offset in the SAR image.

As an extension to this method, the decorrelation filter was used as a pre-processor to enhance the edge content of both images. Then the MNF transform was applied to the SAR image as described above, followed by normalized correlation between this remapped SAR chip and the optical chip.

The results of the MNF method proved promising, and the additional prior use of the decorrelation filter increased the correlation coefficient in the match surface in most cases (see Table 3-2).

3.3.2 Contour-Based Matching Results

The intent in using contour-based matching was to utilize long, linear features of interest, such as roads or power lines. Such features are more likely to be highly visible in the optical images, which generally have a higher signal to noise ratio and often are higher resolution.

Once these linear features have been segmented from the EO imagery, they are used to cue a variety of segmentation and matching methods for the SAR image. Also, since these features form image contours, they can be used to estimate the residual rotation between the SAR and optical images. Therefore, contour-based matching can be used as a step for removing such residual rotation prior to area-based matching, which are especially sensitive to such rotational errors.

The variable profiling algorithm was tested on several Landsat images. It was found that due to the relatively low noise content of the Landsat data, it was possible to use a relatively small profile sum region. As a result, profiling in the vertical direction only provided sufficient filtering to extract contours up to 45° with respect to the horizontal. A complementary profiling in the horizontal direction should suffice to reliably extract all contours of interest.

The width of the region over which the profile is summed determines both the level of noise tolerance and, as well as the angle of the line that can be detected. Wider profiles achieve greater noise immunity but limit the angle of the detectable line. This is the reason for using the profiler first on the lower noise optical imagery to provide cues for searching in the higher noise SAR imagery. Once the direction of a line is known, the profiler can be customized for direction in the high noise imagery so that a wider profile can be used.

The results of the variable profiler on the Landsat band-4 image of the Yuma scene is shown in Fig. 3-6. Once such an image has been created, unconnected segments are removed and adjacent segments are connected together into streaks which are stored as linked lists. These lists can then be accessed according to length for use in cueing potential corresponding locations in the SAR imagery.

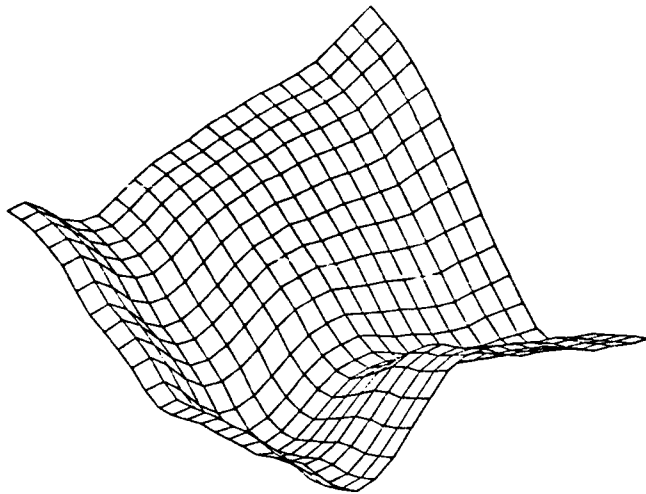
These lists could also be used in image-map matching, assuming the map data were in a digital form such as DMA's DFAD. The results of the variable profiling method for streaks of length 4 segments is shown in Fig. 3-7. Similarly the results for streak lengths longer than 8 segments is shown in Fig. 3-8.

Once potential streaks have been identified as possible features to be identified in the SAR imagery, a variety of algorithms can be used to segment them in the SAR image. Following streak segmentation, image coordinates are used for resampling.

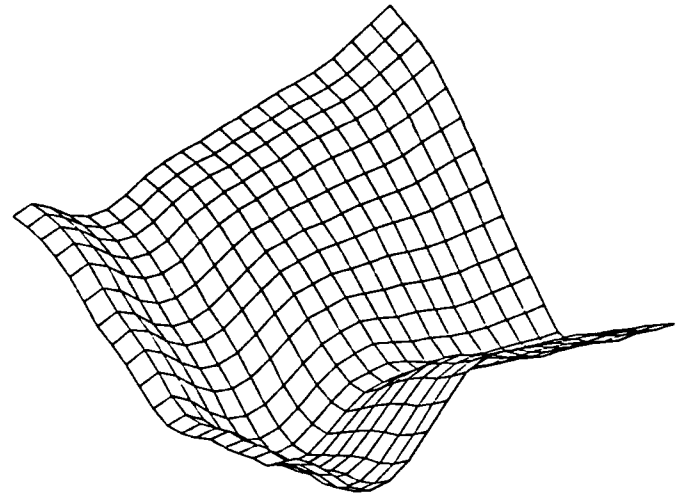
Given a streak and its extent in the optical image, mission parameters define a region of uncertainty in the SAR image within which the corresponding feature must exist. In the worst case, this region will contain the entire SAR image. More likely, however, the uncertainty region involves only a smaller subset of the SAR image.

Within this uncertainty region, a "total profile" is created. The total profile consists of a surface generated by multiple summed profiles at successive trial offsets within this region. At each trial offset, a set of profiles corresponding to the shape of the streak extracted from the optical image is summed, as described in section 3.2.2.1. Fig. 3-9 shows two total profiles generated from the SEASAT SAR of the Yuma site, and the corresponding profiles superimposed with the original SAR images.

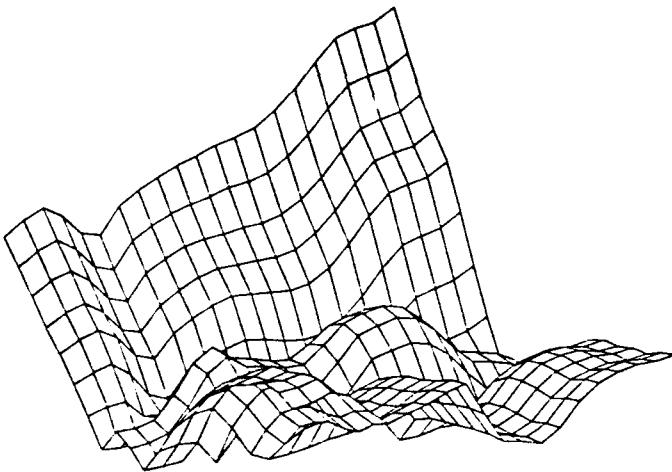
Each offset in the total profile corresponds to pixel in the SAR image. Analysis of these profiles and comparison with the corresponding SAR and optical image pair shows clearly that the total profile method produces a surface containing a peak or ridge along the exact location of the target streak in the SAR image. This peak is especially pronounced for the river and road features in the Yuma site.



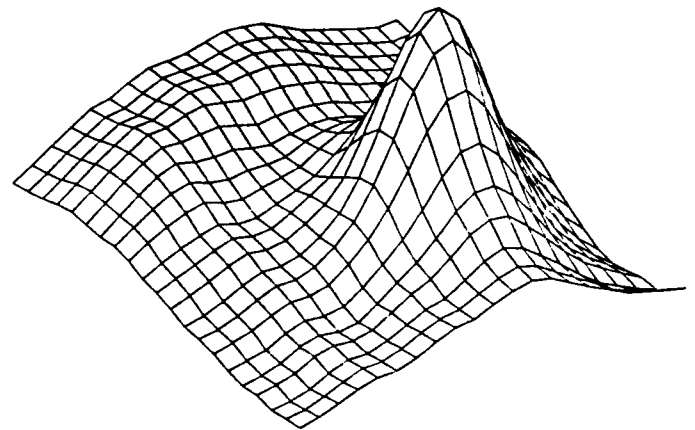
Normalized Correlation Test # 7



Histogram Remap Test # 7



Edge Matching Test # 7



K-L Remap Test # 7

Fig. 3-5: Correlation surfaces derived from normalized correlation of indicated remapped versions of the Landsat Band 4 and SEASAT image pairs of Altamaha River GA site.



Fig. 3-6: Results of the variable profiler on the TM Band 4 of Yuma before connection into linked lists.

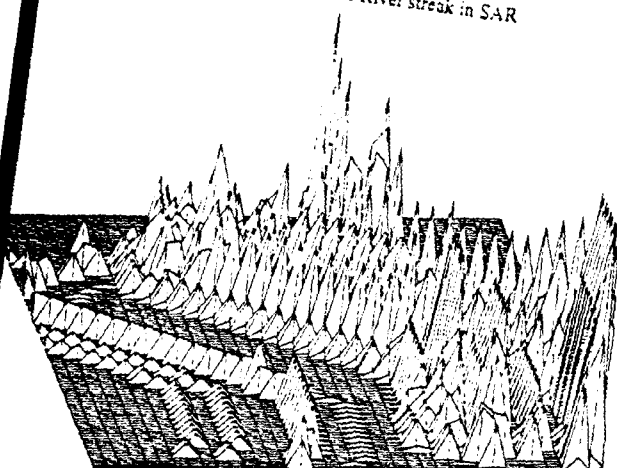


Fig. 3-7: Linked streaks in TM Band 4 of Yuma that are 4 or more segments in length.



Fig. 3-8: Linked streaks in TM Band 4 of Yuma that are 8 or more segments in length.

Summed trial offsets of River streak in SAR



Summed trial offsets of Road streak in SAR

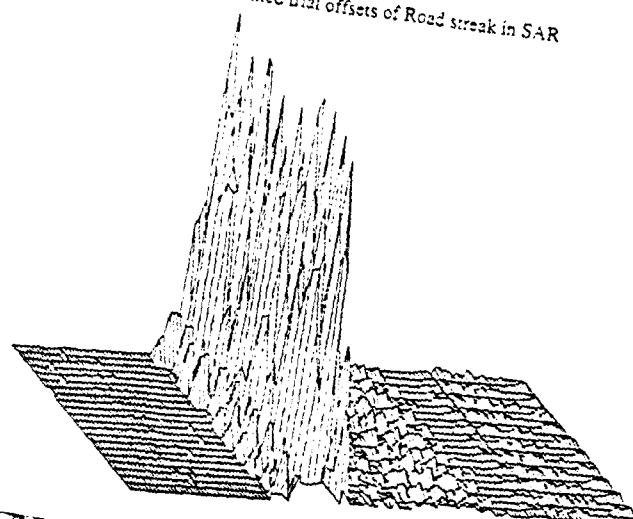
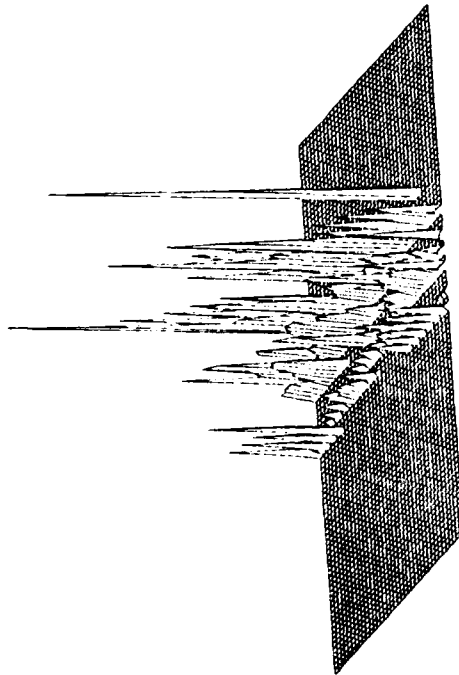


Fig. 3-9: Total profiles of (a) river streak, (b) road streak derived from SEASAT SAR of Yuma region, (c) total profiles shown superimposed on the image with river streak in top center, and road streak in lower right. (Note: In Table 3-2, Road 1 refers to road at left center, Road 2 refers to horizontal road at lower right, and railroad in diagonal lines at lower left.)

Correlation surf of River streak in SAR



Correlation surf of Road streak in SAR

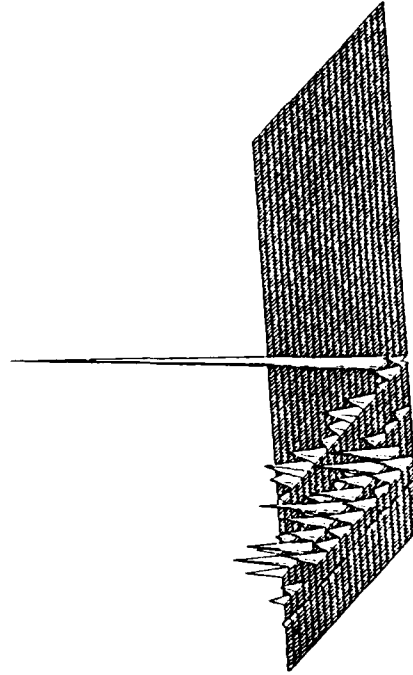


Fig. 3-10: Correlation surfaces derived from sum-of-absolute-differences metric applied between optical streaks and total profile.

Table 3-2 Numerical results of contour-based registration.

Feature	Location in Optical Imagery	Expected SAR Offset	Best SAR Offset	Error (pixels)
River	199, 192	200, 192	204,192	4
Road 1	7, 300	7, 302	7,301	1
Road 2	261, 430	266,427	266,425	2
Railroad (diag)	188, 426	188,423	189,423	1

Table 3-3 Numerical results for combined contour-area-based registration.

Test	Location in Optical Imagery	Expected SAR Offset	Best SAR Offset	Correlation Coefficient
#5 K-L	370, 430	358, 421	358, 422	.8775
#5 MNF	370, 430	358, 421	370, 420	.9308
#6 K-L	285, 338	273, 229	293, 231	.7431
#6 MNF	285, 338	273, 229	271, 223	.9758

The streak first extracted from the optical image then acts as a mask for the a sum of absolute differences correlation (see section 3.1.2). Since the total profile surface has its maximum peaks along a shape which mirrors the feature's imaged shape, the correlation metric should yield the correct feature offset within the SAR image. The resulting correlation surfaces using the sum of absolute differences metric for comparing the total profile and the optical streak mask are shown in Fig. 3-6, with numerical results tabulated in Table 3-2.

As can be seen in Fig. 3-10, the correlation method produces distinct peaks at the correct offset location of the start of the optical profile in the SAR image. The surface given in Fig. 3-7a is slightly misleading in that the peak in the foreground is actually much larger than the peak in the background, but are similar looking because of the perspective distortion of the surface. The road feature shown in Fig. 3-9 and 3-10 refers to ROAD 2 in Table 3-2.

As an enhancement, the profiling in the SAR image can be enlarged in the direction along the candidate streak so that the profile is more robust to speckle but still sensitive to the required shape.

Several other algorithms for segmenting cued features from SAR imagery are documented in [Adair et al,90]. The methods discussed include Randomness Center of Mass, Kolmogorov-Smirnov Test, difference Edge Detector, and Mean-Squared to Variance Ratio test. These methods will be investigated in Phase II.

The correspondences found by such methods produce a set of "tie points" which are then interpolated to form a resampling grid (see section 4.2).

3.3.3 Combined Area-Contour Matching Results

It has been found that successful use of the correlation methods described in section 3.3.1 strongly depends on an initial rotational alignment of the image pair. The remapping and correlation algorithms are quite sensitive to residual rotational errors.

Contour-based methods, however, can be used effectively to estimate this residual rotation between corresponding vectorized contours in the two images at various locations. Following a global rotation, one of the images is then resampled to match the other. This process is then followed by an area-based matching process.

The streak algorithm was applied on the SAR and TM imagery of the Yuma site, and a residual rotation between them of 1.73° was estimated. This rotational error was corrected using a simple bilinear interpolation on the TM image.

Tests #5 and #6 (see Table 3-1) were performed again using the K-L and MNF transformations. A comparison of Tables 3-1 and 3-2 shows a definite improvement in the correlation coefficient of the results for the MNF method. However little improvement can be noted in the actual offset

determined by the methods, except in the case of the K-L which exhibits marked improvement from an error of 12 pixels to an error of 1 pixel with high correlation of about .88.

3.4 Remaining Problems

Considerably more testing and evaluation of test results over a wider variety of sensing scenarios must be performed to better understand qualitatively the circumstances under which the area-based and contour-based registration algorithms are successful or fail. Once such a qualitative understanding is achieved, a quantitative model for the registration process can be attempted.

A goal for Phase II should be to develop a performance prediction model for the success of registering an image pair by a given method, given quantitative sensor parameters as inputs.

4. Sub-Pixel Registration Estimation

In this section is a discussion of two sub-pixel translation estimation algorithms. Such local estimates are then interpolated in order to obtain a global deformation grid valid for the entire image. To support pixel-level change detection then requires the application of a resampling operation.

After initial matching to the nearest pixel has been accomplished, often what is done is to apply some type of interpolation function of the matching function surface in the neighborhood of the correct offset. This serves to locate the position of the extreme point of the match surface to sub-pixel precision. The choice of appropriate interpolation function is an issue.

However, examining this problem from a signal processing point of view, the phase function of a frequency component in the Fourier Transform of an image is much more sensitive to misalignments than the corresponding magnitude function. This observation is the basis for most of the recent work in sub-pixel registration.

4.1 Estimation of Local Translation

The preferential use of the Fourier phase is the basis of the two algorithms below. Algorithm #1 exploits the Fourier Shift Theorem for sub-pixel offset values by differencing phases for corresponding frequencies. Algorithm #2 uses a sophisticated interpolation of the phase correlation surface.

4.1.1 Algorithm #1

The following is a 1-D version of an algorithm for sub-pixel registration estimation. It estimates rational sub-pixel offsets using analytic relationships among non-integer shifts involving the Discrete Fourier Transform (DFT).

One would like to exploit the linear phase relationship for arbitrary translations that exists for the continuous function version of the Fourier Shift Theorem [Papoulis,77].

$$f(t-a) \leftrightarrow e^{-ja\omega} F(\omega)$$

where f, F indicate the corresponding time-domain and Fourier domain functions.

Briefly, the 1-D version is as follows:

- o Apply the FFT to each row separately in both windows. For example, the i^{th} row in the window of images 1 and 2 would be:

$$(a_1, \dots, a_N)_{(i)} = \text{FFT}\{(x_1, \dots, x_N)_{(i)}^{(1)}\}$$

$$(b_1, \dots, b_N)_{(i)} = \text{FFT}\{(x_1, \dots, x_N)_{(i)}^{(2)}\}$$

Do for all rows i , $i = 1, \dots, M$, in both images.

- o For each pair $\{a_k^i, b_k^i\}$ in each row i , with k corresponding to the frequency index, where

$$a_k^i \in (a_1, \dots, a_N)_{(i)}, \quad b_k^i \in (b_1, \dots, b_N)_{(i)}$$

form:

$$I(a)_k^i = \text{imag} \{a_k^i\}, \quad I(b)_k^i = \text{imag} \{b_k^i\}$$

$$R(a)_k^i = \text{real} \{a_k^i\}, \quad R(b)_k^i = \text{real} \{b_k^i\}$$

$$\phi(a)_k^i = \tan^{-1} (I(a)_k^i / R(a)_k^i)$$

$$\phi(b)_k^i = \tan^{-1} (I(b)_k^i / R(b)_k^i)$$

$$d_k^i = [\phi(a)_k^i - \phi(b)_k^i] / (2\pi k / N)$$

$$d_i^* = (1/N) \sum_{k=1}^N d_k^{(i)}$$

- o Cluster the $\{d_i^*\}$. The values of the dominant cluster represent the horizontal shift value. These clustered values should be averaged or used in some estimation procedure.

Only the 1-D version of this algorithm has been tested. The extension to 2-D is expected to occur in Phase II. An example of the performance of this procedure as a function of frequency is shown in Fig. 4-1.

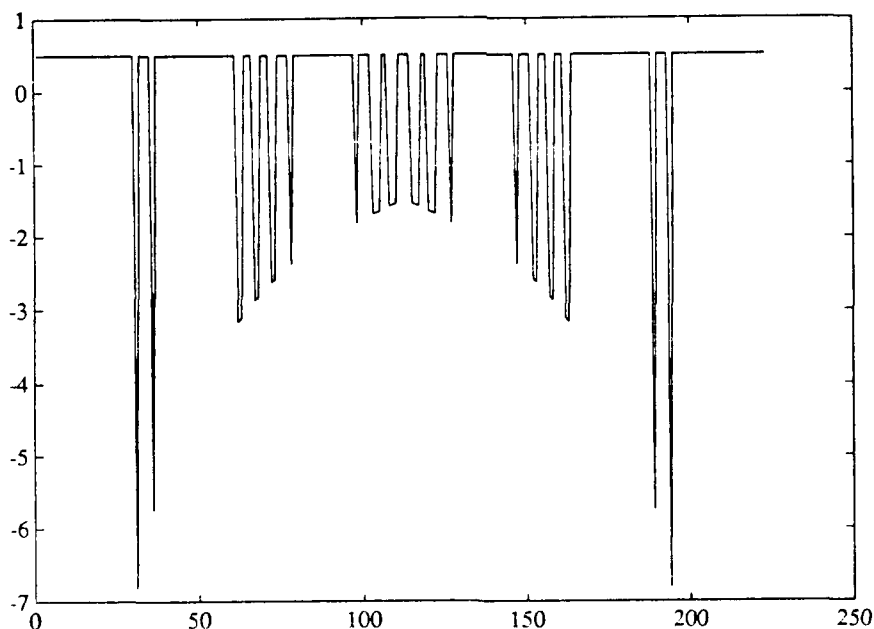


Fig. 4-1: Behavior of Phase for Sub-Pixel Offsets

Some additional theoretical considerations for the above algorithm follow below. In particular, the effects of trying to estimate a sub-pixel offset for two sampled images without prior low-pass filtering are discussed.

Suppose there are two 1-D time sequences, $g(k)$ and $h(k)$. Suppose further that $h(k)$ represents a sub-pixel shifted version of $g(k)$. The following represents a model which is appropriate for estimating this sub-pixel offset.

Let $f(t)$ be a continuous time, finite energy signal and let $F(\omega)$ be the corresponding continuous Fourier transform. Then let $g(k)$ be a sampled version of $f(t)$.

Therefore:

$$g(k) = f(kt_0) \text{ where } t_0 \text{ is the sampling frequency.}$$

$$h(k) = f(kt_0 - a)$$

Since $g(k)$ and $h(k)$ are both discrete time, finite energy sequences, the appropriate version of the Fourier transform is the Discrete Time Fourier Transform (DTFT). We denote these correspondences between domains as:

$$g(k) \xleftrightarrow{\text{DTFT}} G(\exp[j\omega t_0])$$

$$h(k) \xleftrightarrow{\text{DTFT}} H(\exp[j\omega t_0])$$

The DTFT and its inverse are described in [Roberts and Mullis,87], ie:

$$g(k) = (t_0/2\pi) \int_0^{\omega_0} G(\exp[j\omega t_0]) \exp[j\omega k t_0] d\omega$$

$$G(\exp[j\omega t_0]) = \sum_{k=-\infty}^{\infty} g(k) \exp[-jk\omega t_0]$$

where $\omega_0 = 2\pi/t_0$

The characterization of $G(\exp[j\omega t_0])$ and $H(\exp[j\omega t_0])$ in terms of $F(j\omega)$, the continuous time Fourier transform of the continuous function $f(t)$ is now required.

It is shown in [Roberts and Mullis,87] that:

$$G(\exp[j\omega t_0]) = \sum_{k=-\infty}^{\infty} F(j\omega + jn\omega_0)$$

Therefore $G(\exp[j\omega t_0])$ is represented as an infinite sum of shifted versions of $F(j\omega)$. If $F(j\omega)$ is band-limited to $\omega = \omega_0/2$, then:

$$F(j\omega + n\omega_0) = 0 \text{ for all } n \neq 0 \quad (4-1)$$

This implies:

$$G(\exp[j\omega t_0]) = F(j\omega)$$

If the condition in equ (4-1) is not satisfied, then $G(\exp[j\omega t_0])$ will be an aliased version of $F(j\omega)$.

In order to describe $H(\exp[j\omega t_0])$, we let:

$$h(k) = t_0 f(kt_0 - a)$$

Using the inverse DTFT on $h(k)$ and the inverse Fourier transform on $f(t)$, we obtain:

$$\begin{aligned} & (t_0/2\pi) \int_0^{\omega_0} H(\exp[j\omega t_0]) \exp[j\omega t_0] d\omega \\ &= (t_0/2\pi) \int_{-\infty}^{\infty} F(j\omega) (\exp[j\omega(kt_0 - a)]) d\omega \end{aligned} \quad (4-2)$$

But the right hand side of (4-2) can be expressed as:

$$\begin{aligned} & (t_0/2\pi) \int_{-\infty}^{\infty} F(j\omega) (\exp[j\omega(kt_0 - a)]) d\omega \\ &= (t_0/2\pi) \int_{-\infty}^{\infty} F(j\omega) (\exp[j\omega kt_0]) \exp[-j\omega a] d\omega \end{aligned} \quad (4-3)$$

Breaking the infinite limits of the integral into a sum of finite segments, (4-3) can be expressed as:

$$\begin{aligned} & (t_0/2\pi) \int_{-\infty}^{\infty} F(j\omega) (\exp[j\omega kt_0]) \exp[-j\omega a] d\omega \\ &= (t_0/2\pi) \sum_{n=-\infty}^{\infty} \int_{n\omega_0}^{(n+1)\omega_0} F(j\omega) \exp[j\omega kt_0] \exp[-j\omega a] d\omega \end{aligned} \quad (4-4)$$

Let $\phi = \omega - n\omega_0$, then $d\phi = d\omega$. This change of variables converts (4-4) to:

$$\begin{aligned} & (t_0/2\pi) \int_{-\infty}^{\infty} F(j\omega) (\exp[j\omega kt_0]) \exp[-j\omega a] d\omega \\ &= (t_0/2\pi) \sum_{n=-\infty}^{\infty} \int_0^{\omega_0} \exp[-j(\phi + n\omega_0)a] F(j\phi + jn\omega_0) \exp[jk\phi t_0] \exp[jkn\omega_0 t_0] d\phi \end{aligned} \quad (4-5)$$

But the exponential factor $\exp[jkn\omega_0 t_0] = 1$ since $\omega_0 t_0 = 2\pi$. Therefore, the previous integral expression is given by:

$$(t_0/2\pi) \sum_{n=-\infty}^{\infty} \int_0^{\omega_0} \exp[-j(\phi + n\omega_0)a] F(j\phi + jn\omega_0) \exp[jk\phi t_0] d\phi$$

Therefore, using absolute convergence one can interchange the sum and integral [Fulks, 66]:

$$\begin{aligned} & (t_0/2\pi) \int_0^{\omega_0} H(\exp[j\omega t_0]) \exp[j\omega t_0] d\omega \\ &= (t_0/2\pi) \sum_{n=-\infty}^{\infty} \int_0^{\omega_0} \exp[-j(\phi + n\omega_0)a] F(j\phi + jn\omega_0) \exp[jk\phi t_0] d\phi \\ &= (t_0/2\pi) \int_0^{\omega_0} \sum_{n=-\infty}^{\infty} \exp[-j(\phi + n\omega_0)a] F(j\phi + jn\omega_0) \exp[jk\phi t_0] d\phi \end{aligned} \quad (4-6)$$

Comparing the right-hand side of (4-6) with the left-hand side of (4-2) implies:

$$H(\exp[j\omega t_0]) = \sum_{n=-\infty}^{\infty} F(j\omega + jn\omega_0) \exp[-j(\omega + n\omega_0)a] \quad (4-7)$$

Now, if $a = 0$, then:

$$H(\exp[j\omega t_0]) = G(\exp[j\omega t_0]) = \sum_{n=-\infty}^{\infty} F(j\omega + jn\omega_0) \quad (4-8)$$

Equations (4-7) and (4-8) will now be used to estimate the sub-pixel shift a . Suppose $F(j\omega)$ is band-limited to $\omega_0/2$, then:

$$F(j\omega) = 0 \text{ for } |\omega| > \omega_0/2, \text{ ie. } F(j\omega + jn\omega_0) = 0 \text{ for } n \neq 0.$$

This implies that:

$$G(\exp[j\omega t_0]) = F(j\omega)$$

$$H(\exp[j\omega t_0]) = F(j\omega) \exp[-j\omega a]$$

Using a discrete Fourier transform, one can compute samples of $G(\exp[j\omega t_0])$ and $H(\exp[j\omega t_0])$.

Taking a ratio of the samples of $G(\exp[j\omega t_0])$ and $H(\exp[j\omega t_0])$, one can compute $\exp[j\omega a]$ which gives the sub-pixel shift a .

If $F(j\omega)$ is not band-limited, then $F(j\omega + jn\omega_0)$ is non-zero for non-zero values of n . This results in an aliased spectrum in both cases. As a result of the aliasing, one cannot easily compute the sub-pixel offset.

If $f(t)$ is not a band-limited signal, it can be frequency filtered to make it band-limited. It is critical that the filtering be performed prior to sampling. Otherwise, aliasing will again result.

4.1.2 Algorithm #2

Another local translation estimator which was investigated in Phase I is the phase correlation algorithm [Schaum and McHugh,88], [Stocker and Clayton,89], [Stocker,90]]. Although phase correlation is generally known to be an effective method for measuring the translational misregistration between a pair of optical image sub-frames, there are several "fine points" involved in applying the technique to discrete data. In the following, we review the basic phase correlation approach and its underlying model, discuss several important details of its implementation, show some test results on synthetically-shifted data, and apply the method to measure residual misregistrations in the Death Valley scenes.

Basic Model. Consider a continuous optical background scene $s(x,y)$ that is projected onto the focal plane on Frame 1. At Frame 2, the same scene appears displaced by an arbitrary 2-D spatial offset (α, β) due to apparent background translation. The images corresponding to these two frames are defined as

$$\phi_1(x,y) = s(x,y) \quad (4-9a)$$

$$\phi_2(x,y) = s(x-\alpha, y-\beta) \quad (4-9b)$$

Let $S(u,v)$ be the 2-D Fourier transform of the scene $s(x,y)$, where (u,v) are spatial frequencies corresponding to the focal plane position variables (x,y) . The 2-D Fourier transforms of the two frames are

$$\phi_1(u,v) = S(u,v) \quad (4-10a)$$

$$\phi_2(u,v) = S(u,v)e^{-j2\pi(\alpha u + \beta v)} \quad (4-10b)$$

The 2-D shift manifests itself as an additive linear phase in each dimension of the spatial frequency domain.

The optical images $\phi_1(x,y)$ and $\phi_2(x,y)$ are sampled on the focal plane to generate the discrete sensor frames $f_1(k,l)$ and $f_2(k,l)$. We assume without loss of generality that the scene transform $C(u,v)$ is bandlimited by the sensor optics and detectors to the spatial frequency interval $[-0.5, +0.5]$ in both the u and v dimensions. Then the sampling interval can conveniently be defined as unity in each spatial coordinate. If each frame is space-limited to N pixels in the k and focal plane coordinates and viewed as a periodic function, then it can be shown that the discrete Fourier transforms defined by

$$F_i(m,n) = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f_i(k,l) \exp\{-j2\pi(km+ln)/N\} \quad (4-11)$$

$i=1,2$

provide N samples of the respective frame Fourier transforms $\phi_i(u,v)$ on the spatial frequency repetition interval from $[-0.5, +0.5]$ in both u and v . These samples occur at evenly-spaced frequency increments of $1/N$ in either dimension, so that

$$F_i(m,n) = \phi_i(m/N, n/N) \quad (4-12)$$

$$-N/2 \leq m \leq N/2-1, \quad -N/2 \leq n \leq N/2-1$$

and from 4-(10) we can write

$$F_1(m,n) = S(m/N, n/N) \quad (4-13a)$$

$$F_2(m,n) = S(m/N, n/N) e^{-j2\pi(\alpha m + \beta n)/N} \quad (4-13b)$$

To simplify the notation, we will use $S(m,n)$ to represent $S(m/N, n/N)$ in the subsequent discussion.

Phase Correlation. The phase correlator is a whitened cross-correlation estimate defined by

$$r(\Delta x, \Delta y) = \sum_m \sum_n w(m,n) [F_1^*(m,n) F_2(m,n)] / [|F_1^*(m,n) F_2(m,n)|] \times \exp\{j2\pi(m\Delta x + n\Delta y)/N\}$$

as a function of continuous 2-D lag variables $(\Delta x, \Delta y)$. The spatial frequency weight $w(m,n)$ is included here for generality; its role is discussed later.

For the pair of perfectly translated frames F_1 and F_2 defined in (4-13), the above phase correlation function evidently peaks at the true 2-D offset $(\Delta x, \Delta y) = (\alpha, \beta)$. The translation between the two frames can therefore be determined by precisely measuring the location of the peak of (4-14).

Phase correlation is best viewed as a weighted cross-correlation of two "whitened" image frames whose spatial frequency components are defined by $F_1(m,n)/|F_1(m,n)|$ and $F_2(m,n)/|F_2(m,n)|$, respectively. A whitened frame retains only the phase information in the original spatial frequency components, and is often referred to as a phase-only image. Spectral whitening is a nonlinear filtering process which tends to enhance the high-frequency scene features (such as edges) that contain most of the useful information for registration.

Since many natural background scenes from electro-optical sensors tend to have a substantial amount of power concentrated at lower spatial frequencies, a conventional (unwhitened) cross-correlation often produces a relatively broad output peak. The spectral whitening used by the phase correlation procedure tends to sharpen the peak, leading to more precise measurements of frame displacement.

Discrete Implementation. The phase correlator in (4-14) is theoretically defined on the continuous 2-D spatial variables $(\Delta x, \Delta y)$. However, in a digital computer, the correlation is actually computed at discrete intervals. If the desired lag interval in each dimension is selected as one pixel, then the phase correlation can be computed by the 2-D inverse DFT

$$r(p,q) = \sum_{k=-N/2+1}^{N/2-1} \sum_{l=-N/2+1}^{N/2-1} [x_1^*(m,n) x_2(m,n)] \times \\ 1/[|x_1(m,n) x_2(m,n)|] \times \exp\{j2\pi(pm+qn)/N\} \quad (4-15)$$

for integer-valued pixel lags (p,q) in the range $-N/2, \dots, N/2-1$. Note that the DFT summations are performed symmetrically about the dc spatial frequency component located at $(m,n) = (0,0)$ to ensure a real-valued result.

The phase correlation samples in (4-15) can be found by brute-force calculation or by taking an $N \times N$ inverse FFT of the weighted cross-spectrum phase function. The FFT approach is most efficient unless the required number of lags is known to be fairly small. It provides circular phase correlation samples for pixel lags ranging from $-N/2$ to $N/2-1$ in either spatial dimension. A non-circular implementation can be obtained by zerofilling both frames prior to the forward FFTs, but this requires a larger FFT size. A less-expensive way of avoiding problems due to circular correlation and the non-periodicity of actual frame data is to apply a tapered weighting function (such as a 2-D Hanning window) to the frames prior to the forward FFT in (4-11) that transforms them to the frequency domain.

The frequency-dependent weighting function $w(m,n)$ in (4-15) is used to

emphasize those regions of the cross-spectrum phase function that contain useful phase information. This function is generally chosen to have a low-pass characteristic (e.g., 1-D Hanning or Gaussian window). The reason is that in most sensor data, the highest spatial frequencies tend to be heavily corrupted by noise and aliased components; thus attenuating these components with frequency-selective weighting improves the misregistration estimates.

The location of the peak in the phase correlation (4-15) is a measure of the relative displacement or misregistration between frames 1 and 2. To estimate the displacement to sub-pixel accuracy, the peak position must be accurately interpolated from the finite set of samples. The classical solution to this problem involves the application of the appropriate 2-D "sinc" interpolation function on the discrete cross-correlation estimate. A far more practical approach is to fit a polynomial to the cross-correlation samples in the vicinity of the highest peak. A separable quadratic polynomial is the simplest choice since it uses only three points in either dimension: the peak sample and its neighbors to either side. Although the translation estimates provided by this simplified interpolator are biased, it can be shown that the bias is a function of known processing parameters [Stocker and Clayton, 89], [Stocker, 90]. Thus, it is possible to correct the measurement biases on the fly using pre-computed lookup tables.

A complete block diagram of the phase correlation based registration algorithm, including the bias-corrected peak measurement step, is shown in Figure 4-2. This algorithm has been successfully calibrated on synthetically-shifted data as noted below.

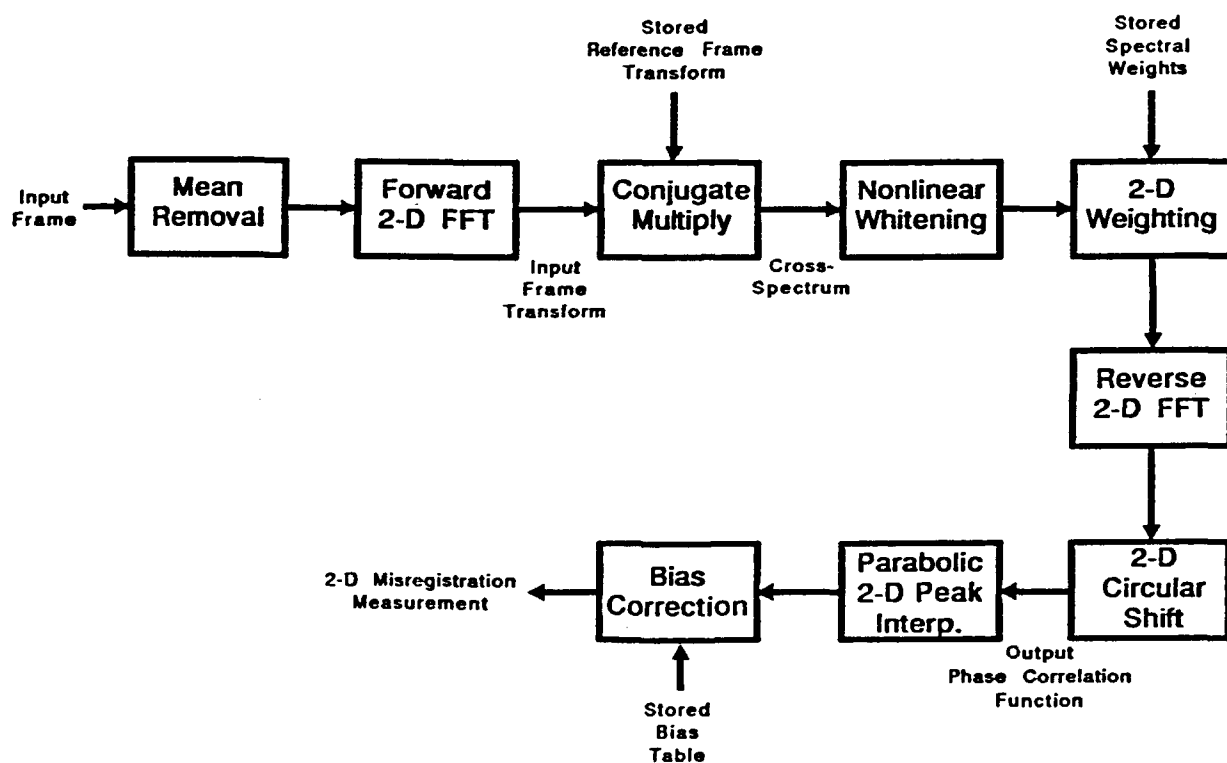
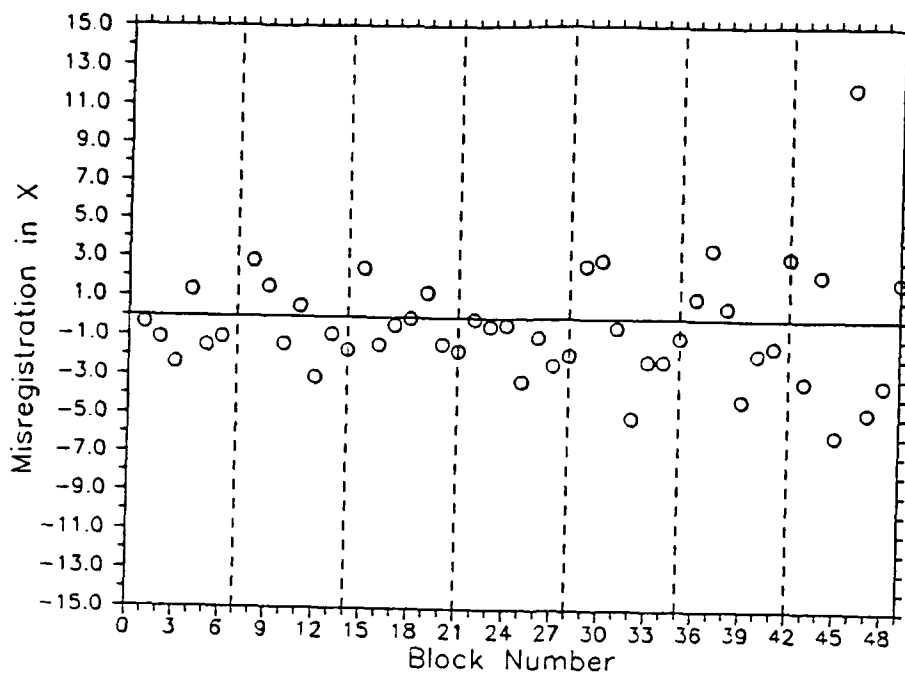
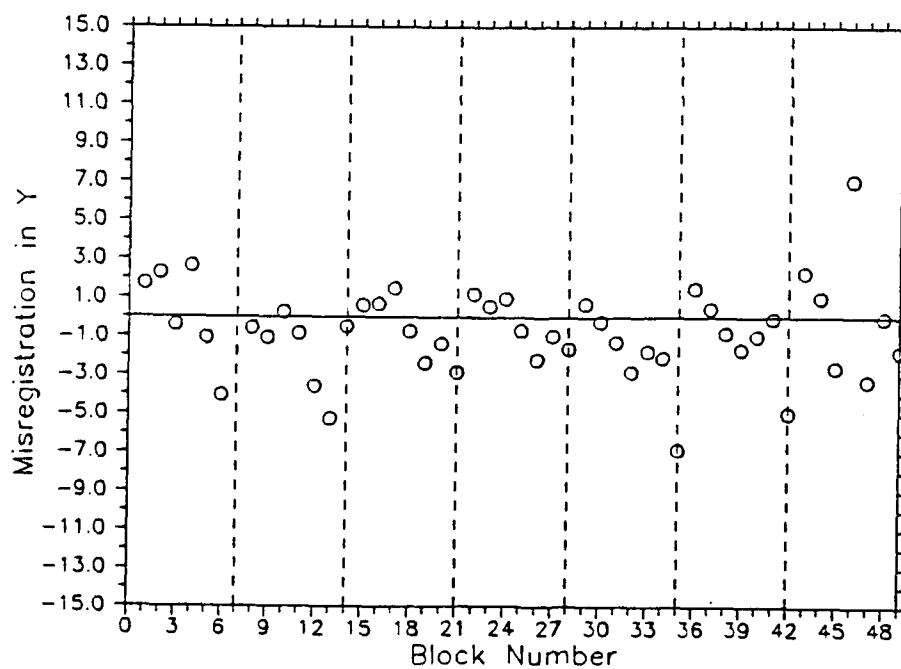


Fig. 4-2: Block diagram of Registration Algorithm Based on the Phase correlation Method



(a)

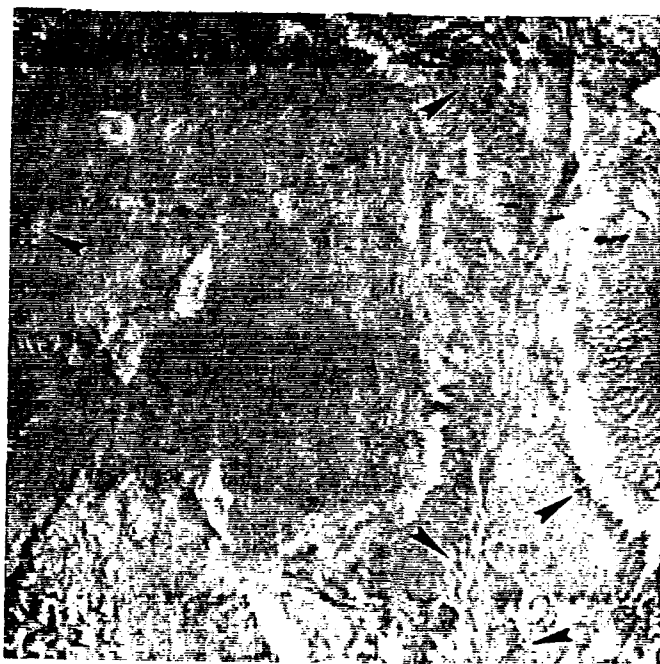


(b)

Fig. 4-3: Correlation Block Translation Measurements for Death Valley Data, (a) horizontal dimension, (b) vertical direction.



(a)



(b)

Fig. 4-4: Weighted Difference Images for Death Valley (1983 and 1988 Scenes) - Band 1, (a) After Global Registration, (b) After Fine Registration.

4.2 Global Interpolation of Local Estimates

Following the computation of offsets for smaller sized blocks of pixels, an interpolation process will be required to attain "subpixel" accuracy.

One problem with global least-squares matching for interpolation of registered control points is that errors are averaged over the entire image. In fact, what is really required is to find an interpolation scheme that is only influenced locally by the data, since the distortions due to terrain, atmospheric effects, and sensor nonlinearities act are spatially varying and so should be compensated in a locally adaptive manner.

Another approach is to use splines, which restricts the effects to be more local. Consider the surface spline representing an infinite plate under the imposition of point loads [Goshtasby,88]:

$$\text{Let: } f(x,y) = a_0 + a_1 x + a_2 y + \sum_{i=1}^n F_i r_i^2 \ln r_i^2$$

where:

$$n = \# \text{ loads}$$

$$r_i^2 = (x-x_i)^2 + (y-y_i)^2$$

$$(x_i, y_i) = \text{position of } i^{\text{th}} \text{ control pt.}$$

$$f(x,y) = \text{elevation of surface at } (x,y)$$

Now, substituting control pts. into equation for f and solving the following set of equations leads to the parameters: $a_0, a_1, a_2, F_i, i=1, \dots, n$.

$$\sum_{i=1}^n F_i = 0$$

$$\sum_{i=1}^n x_i F_i = 0$$

$$\sum_{i=1}^n y_i F_i = 0$$

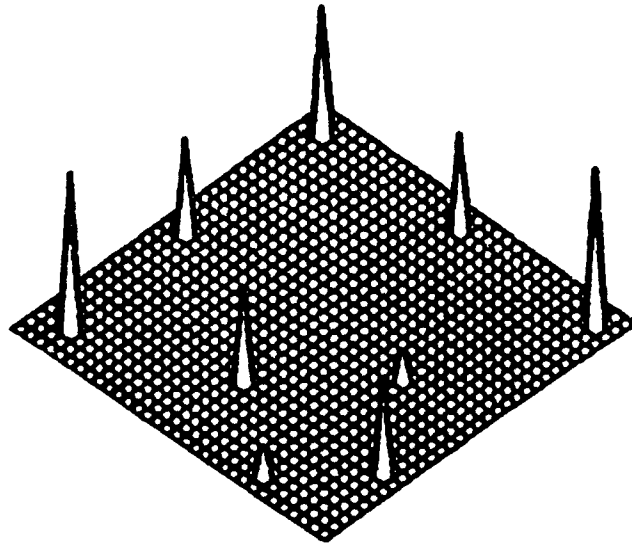
$$f(x_1, y_1) = a_0 + a_1 x_1 + a_2 y_1 + \sum_{i=1}^n F_i r_{i1}^2 \ln(r_{i1}^2)$$

.
.
.

$$f(x_n, y_n) = a_0 + a_1 x_n + a_2 y_n + \sum_{i=1}^n F_i r_{in}^2 \ln(r_{in}^2)$$

This method will be investigated in Phase II for interpolating registration control points.

Translation Measurements (Control Points)



Surface Fit to Measurements (Thin-Plate Interpolating Spline)

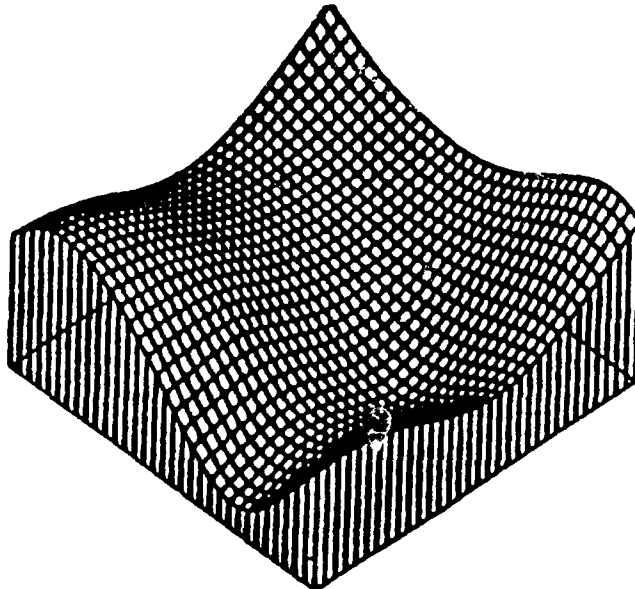


Fig. 4-5: Surface Fit to a Sparse Set of Translation Measurements

A related global interpolation method was implemented and tested during Phase I. This method uses the thin-plate cubic spline [Lancaster and Salkauskas,86]. Physically, the thin plate is analogous to a mechanical surface spline that is not subjected to torques at the data points and deforms only by bending. The thin plate spline is mathematically formulated as the 2-d equivalent of the well-known natural cubic spline, which is optimal in the sense minimizing total stress energy.

The thin plate is formally derived from a Boolean sum of two different projection operators:

- a) An interpolating projection formed from translates of the basis function $r_i^2 \ln(r_i)$ as defined above,
- b) A weighted least-squares projection with a weight matrix equal to the inverse of the Vandermonde matrix of the translates of the same basis function.

It should be noted that the computation of this spline can be computationally relatively expensive, since it involves the inversion of a matrix of dimension N , where N is the number of data points to be interpolated.

An example of thin-plate spline interpolation applied to the generation of a global deformation surface is shown in Fig. 4-5. A set of 2-D translation measurements using control points taken at a discrete set of locations within the image frame is created. Then a thin-plate interpolating spline is computed for the 2-D apparent translation vs. pixel position. These surfaces are then used for image resampling.

This technique was successfully applied to to perform global registration for the Death Valley IR imagery discussed in section 2.4.

4.3 Image Resampling

Accurate resampling of the discrete image frames is critical to the performance of change detection schemes based on weighted frame differencing. Even if the global registration error measurements are perfect, a poorly implemented resampling step can introduce significant noise into a difference image.

The use of sliding-window FIR interpolators is a viable approach to the image resampling problem, provided that a suitable interpolation kernel is selected [Wolberg,90]. In this approach, a discrete image is resampled by means of a 2-D convolution operation. Assume that global registration calls for the image frame $I(m,n)$ to be shifted by a variable amount $(\Delta x_{mn}, \Delta y_{mn})$ as a function of the pixel location (m,n) . Let

$$(\Delta x_{mn}, \Delta y_{mn}) = (p_{mn}, q_{mn}) + (\alpha_{mn}, \beta_{mn}),$$

where (p,q) and (α,β) denote the interger and fractional parts of the desired pixel shifts, respectively. The resampling operation produces a

new discrete frame I_r defined by

$$I_r(m+p_{mn}, n+q_{mn}) = \sum_k \sum_l x(m+k, n+l) h(k-\alpha_{mn}, l-\beta_{mn})$$

where $h(x,y)$ is the 2-D interpolation kernel. In the present study, only separable kernels of the form $h(x,y) = h(x)h(y)$ with a finite support of N pixel intervals in either dimension are considered.

The performance of various kernels can be interpreted and compared by considering the classical digital signal processing approach to interpolation [Lucke, undated]. In this approach, we convolve the discrete frame data with a continuous (or, in practice, a highly oversampled) reconstruction kernel of finite extent; then decimate the resulting image to obtain interpolated samples on the desired frame grid. Although these steps are often combined into one operation (as in the above equation), it is useful to think of them as two distinct conceptual stages.

The optimum interpolation kernel for sampled bandlimited data is the "sinc" function $\sin(\pi n)/(\pi n)$ with nulls spaced at the discrete pixel intervals indexed $n = +1, +2, \dots, +\infty$. The frequency response of this kernel is constant over the normalized spatial frequency interval $(-0.5, +0.5)$ and zero outside of this interval. It achieves perfect signal reconstruction since it produces no in-band spectral distortion and completely removes the aliases of the sampled signal which repeat indefinitely at the pixel rate. Unfortunately, ideal interpolation can only be approximated on a finite-length data record and, in any case, is too computationally demanding to consider for most applications. Thus in practice, relatively small kernels with frequency responses that approximate the ideal "rect" function are utilized.

The desired characteristics of an interpolation kernel are linear phase, a nearly flat in-band response and very low out-of-band sidelobes. A reasonably flat linear-phase passband reduces the signal distortion due to interpolation filtering. Low sidelobes minimize the amount of interpolation "noise" introduced by aliased signal components that fold into the original band when the filtered signals are decimated (or resampled) at or near the original sampling rate.

Ideally, it would be desirable to apply the same in-band filtering to each image frame to be differenced or otherwise combined in a change detection operation. This condition can be satisfied exactly if each frame is resampled (i.e., filtered and decimated) in precisely the same way, but this generally cannot be done in non-trivial registration applications. The problem is that the different resamplings needed for frame registration apply different phase shifts to the out-of-band components of the interpolated scene. These components, which fold back into the original band upon decimation, constitute a source of "noise" which varies in a rather unpredictable way from one image frame to another. Significant clutter leakage can result due to mismatches in the spatial frequency content of the interpolated frames being differenced, even in cases where the resampling itself is geometrically perfect.

A practical solution to this problem is to choose an interpolation kernel

that attenuates the out-of-band signal components to the point where they are well below the level of the sensor noise. For a typical electro-optical sensor that provides a background clutter-to-noise ratio on the order of (say) 10 (or 20 dB), this implies the use of FIR filter kernels with peak sidelobes on the order of 25-30 dB or better.

To gain some insight into the performance of various FIR interpolators, we evaluated frequency responses and some simple figures-of-merit for several filter kernels. Each kernel was oversampled by a factor of 10 in a symmetric fashion to ensure that its frequency response would have linear phase. The squared-magnitude of the frequency response was then calculated from the DFT formula. Two figures of merit were also computed for each kernel, based on an assumed work-case ac background "signal" having a triangular-shaped amplitude spectrum extending out to the band edges at the normalized spatial frequencies (-0.5, +0.5). These figures-of-merit are defined as follows:

- 1) Signal Loss. The loss in signal power in-band due to the filtering applied by the interpolation kernel (a measure of the amount of passband attenuation).
- 2) Signal-to-Noise Ratio (SNR). The ratio of in-band signal power after filtering by the interpolation kernel (a measure of the relative strength of out-of-band interpolation noise components).

The following interpolation filter kernels were evaluated in this manner:

- a) 4-point cubic convolution
- b) 4-point cubic B-spline
- c) 4-point custom FIR filter
- d) 4-point DFT filter
- e) 6-point DFT filter
- f) 8-point DFT filter

Cubic convolution is a 4-point polynomial kernel originally derived as an efficient approximation to the sinc function [Wolberg,90]. The cubic B-spline kernel is the filter defined by three successive self-convolutions of the unit rectangle function [Wolberg,90]. The custom FIR filter is a special kernel designed by Space Computer Corporation using the Remez Exchange optimization algorithm with special zero placement. The DFT filters, originally derived in close-form by R. Lucke at NRL [Lucke,undated], are obtained by computing the inverse DFT of the N-point linear phase filter in the spatial frequency domain that implements a specified image-domain translation.

Fig. 9-2 through 9-7 (in Appendix 9.3) plot the impulse response and frequency response for each of the above kernels. The calculated figures-of-merit are given in Table 4-1 below.

Kernel	Signal Loss (dB)	SNR (dB)
Cubic Convolution	0.202	24.63
Cubic B-Spline	1.110	37.33
Custom FIR Filter	0.730	33.83
DFT-4 Filter	0.004	26.98
DFT-6 Filter	0.008	31.86
DFT-8 Filter	0.000	35.42

Table 4-1: Figures-of-Merit for Interpolation Kernels

Basic tradeoffs among the kernels are evident. For example, cubic convolution has a good passband characteristic but relatively poor sidelobe response. In contrast, the cubic spline kernel achieves excellent sidelobes at the expense of increased in-band attenuation. The custom 4-point FIR filter represents a somewhat better balance between the conflicting desires for a flat passband and low sidelobes. The even-sample DFT kernels are also interesting in that they combine a nearly perfect passband with reasonably good sidelobe levels.

The performance differences between the three DFT filters illustrate the general trend toward improved response with increasing kernel length or number of coefficients. Of course, a resampling algorithm based on a longer kernel is also relatively expensive to implement. Unless the signal-to-noise ratio is unusually high, it appears that a well-chosen 4-point kernel should suffice for most change detection applications with real imagery. The registration experiments conducted during Phase I utilized resampling algorithms based on either the cubic B-spline or the DFT-4 kernel.

In classical interpolation lore, the use of a filter with a smooth passband response is considered to be of paramount importance to minimize spectral amplitude distortion of the "signal" being interpolated. However, for frame difference signal processing applications, it seems that this criterion might be considerably less important than the out-of-band sidelobe level. For example, how important is it to faithfully reproduce the spectral content of the unchanged portion of the background it is going to be subtracted out later? Is it necessarily bad to attenuate higher in-band frequencies if those frequencies are dominated by sensor noise and/or aliasing components rather than by baseband signal components.

One way to clarify these trades for electro-optical sensor images would be to formulate a composite signal model consisting of "change" signal, a random background, and additive noise; properly accounting for important sensor effects such as optical blur, detector spatial filtering and 2-D focal plane sampling. An interpolation filter response could then be applied to each of these scene components and used to derive more meaningful figures-of-merit, such as the effective signal-to-clutter +

noise ratio after a frame differencing operation. Such an analysis could be pursued under a Phase II program to help guide the final selection of the image resampling algorithm.

4.4 Results

4.4.1 Results for Sub-Pixel Algorithm #1

Preliminary 1-D simulation of a square wave function offset to .3 and .9 pixels was performed for VEXCEL's Algorithm #1. No noise was added to the square wave as yet. The resulting errors for estimation were:

- o 4×10^{-4} pixel for .9 offset,
- o 5×10^{-4} pixel for .3 offset.

4.4.2 Results for Sub-Pixel Algorithm #2

Both extensive simulation and real data were employed for testing SCC's Algorithm #2.

4.4.2.1 Simulation Experiments

Simulation experiments were performed to verify the accuracy of the phase correlation registration algorithm with the above peak interpolation scheme. The experiment made use of synthetically-jittered frames (without added noise) to provide a controlled test of measurement quality for several variations of the basic algorithm.

The experiments were based on a typical long-wave IR scene collected by the MIT Lincoln Laboratory IRMS sensor. To determine the accuracy of the sub-pixel translation measurements provided by the registration algorithm, we used this scene to create a synthetic sequence of three new frames having known displacements in the elevation (or vertical) direction. This was done by Fourier transforming the original 256 x 256 IRMS scene, bandlimiting the transform with a 2-D Hanning spectral window, and zerofilling the windowed array to a length of 1024 points in the vertical direction. The resulting 256 x 1024 array was then transformed back to the image domain to obtain a filtered scene oversampled by a factor of 4:1 in elevation. This scene was then vertically decimated by a factor of 4, starting from three different rows 1, 2 and 3, to produce a set of three new 256 x 256 frames. Adopting the first of these new frames as the registration reference, the second and third frames have respective vertical displacements of precisely -0.25 and -0.50 pixel.

As a check the accuracy of the phase correlation algorithm, we measured the know 2-D shifts between the synthetic frame pairs 2 and 1 and 3 and 1. The measurements obtained from several variations of correlation processing combined with parabolic peak interpolation are summarized in Tables 4-2 and 4-3 below. "Circular Correlation" refers to a conventional (unwhitened) cross-correlation; it produces biased measurements in either case. In contrast, the circular phase correlation measurements, after

appropriate bias correction, are nearly perfect for both the Uniform (i.e., rectangular) spectral window and a heavily tapered Hanning² spectral window function defined by

$$w(m,n) = w(m)w(n) = [1+\cos(2\pi m/N)]^2 [[1+\cos(2\pi n/N)]^2$$

The calibration results given here establish the systematic accuracy of the registration algorithm based on discrete phase correlation combined with an appropriate peak measurement procedure.

4.4.2.2 Application to Real Data

Although phase correlation can be a remarkably accurate procedure for measuring the sub-pixel displacement between a pair of well-correlated image frames, its ability to register moderately correlated imagery with long-term changes is less certain. To gain insight into its performance in this situation, we processed the three-band Death Valley infrared data described in Section 2.4.

The basic assumption was that the residual registration error between the globally-registered 1983 and 1988 Death Valley images could be approximated by a slowly-varying 2-D translation. Local misregistration measurements were obtained by applying the phase correlation procedure described above to 128x128 pixel sub-blocks overlapped by 50% in either dimension. For the 512x512 scene, this resulted in a total of 49 2-D displacement measurements spaced at 64-pixel intervals within the frame.

Since the Death Valley scene appeared to exhibit more long-term correspondence in its principle spectral components than in the original spectral bands, the block phase correlation procedure was separately applied to each of the three spectral components of the 1983 and 1988 multi-band images. For each block, the component which produced the highest correlation peak was used to measure the 2-D translation of the 1988 image with respect to the 1983 reference image. In this way, the spectral component with the best local feature contrast was always used to estimate the local misregistration.

Figure 4-3 plots the local translation measurements in each dimension of the Death Valley scene vs. the correlation block number. For this purpose, the blocks are assigned numbers in standard raster fashion starting in the upper right-hand corner of the scene. Thus correlation block 1 is located in the upper right-hand (or northwest) corner, while block 49 is at the lower left-hand (or southwest) corner. Vertical dotted lines are drawn at 7-block intervals to indicate measurements taken along the same horizontal row of the image. These measurements indicate the presence of more or less random multi-pixel registration errors between the Death Valley scenes following coarse global registration based on manual control points.

There is no obvious way to determine the accuracy of these displacement measurements in the absence of ground truth information. However, one reasonable check is to perform a second fine registration step based on the measurements, and determine whether the pixel-to-pixel correspondence between the two image sets is improved.

Table 4-2 Measurement Results for Synthetically Jittered IRMS Frames 1 and 2.
Actual Jitter in (Azimuth,Elevation) = (0,-0.25) Pixel

Cross-Correlation Processing	FFT Size	Cross-Spectrum Weighting	2-D Jitter Measurement	2-D Measurement Error
Circular Correlation	256x256	Uniform	(-0.024,-0.236)	(-0.024, 0.014)
Circular Phase Correlation	256x256	Uniform	(0.000,-0.252) *	(0.000,-0.002)
Circular Phase Correlation	256x256	Hanning ²	(0.000,-0.250) *	(0.000, 0.000)
Non-Circular Phase Correlation	512x512	Uniform	(-0.001,-0.234) *	(-0.001, 0.016)
Non-Circular Phase Correlation	512x512	Hanning ²	(0.000,-0.247) *	(0.000, 0.003)

* — bias-corrected measurements

Table 4-3 Measurement Results for Synthetically Jittered IRMS Frames 1 and 3.
Actual Jitter in (Azimuth,Elevation) = (0,-0.50) Pixel

Cross-Correlation Processing	FFT Size	Cross-Spectrum Weighting	2-D Jitter Measurement	2-D Measurement Error
Circular Correlation	256x256	Uniform	(-0.048,-0.499)	(-0.048, 0.001)
Circular Phase Correlation	256x256	Uniform	(0.000,-0.500) *	(0.000, 0.000)
Circular Phase Correlation	256x256	Hanning ²	(0.000,-0.500) *	(0.000, 0.000)
Non-Circular Phase Correlation	512x512	Uniform	(-0.002,-0.445) *	(-0.002, 0.055)
Non-Circular Phase Correlation	512x512	Hanning ²	(-0.001,-0.491) *	(-0.001, 0.009)

* — bias-corrected measurements

To do this, the translation measurements made in each dimension were fit to cubic thin-plate splines to obtain smooth 2-D surfaces of displacement vs. pixel position for the 1988 image. These displacement surfaces were then used by a sliding-window interpolator to resample the globally-registered images from each of the three wavebands. The interpolator employed for this purpose was the cubic spline kernel discussed in Section 4.3.

Two basic figures of merit were used to assess whether any improvement was obtained from fine registration. The first was the global average pixel correlation coefficient between corresponding bands of registered imagery, which should increase if the unchanged background component is better aligned. Table 4-4 shows that small increases in the average pixel correlation were indeed observed in each of the three bands due after fine registration.

TIMS Band	Pixel Correlation Coefficient	
	After Global Registration	After Fine Registration
1	0.57	0.60
3	0.59	0.62
5	0.69	0.72

Table 4-4. Pixel Correlation in Death Valley 1983 and 1988 Images

A second indicator of registration performance is the rms level of a minimum-power weighted difference frame formed from each registered band pair, which is affected by clutter leakage due to background misalignment as well as actual long-term changes in the scene. Table 4-5 shows that small improvements in this measure are also observed after the fine registration step.

TIMS Band	RMS Level of Difference Image ($\mu W/sr/cm^2/\mu m$)	
	After Global Registration	After Fine Registration
1	39.4	38.2
2	46.1	44.5
3	37.7	36.3

Table 4-5: RMS Level of Weighted Difference Image for Death Valley 1983 and 1988 Scenes

Figure 4-4 compares the difference images from TIMS Band 1 on the same grey-scale before and after the fine registration based on blockwise phase correlation. These difference images are qualitatively similar in most

respects. However, in several places denoted by black arrows, it appears that the fine registration may have reduced the amount of contrast edge leakage. Most of the pixel-level "changes" observed in these difference frames result from pixel radiance variations caused by local temperature changes between looks.

It appears that the phase correlation technique has the potential to improve the local registration of multi-band imagery to be utilized for change detection. However, it also seems clear that the effectiveness of this algorithm (and all other area-based correlation methods) will be reduced by image-to-image decorrelation which results from long-term pixel intensity fluctuations.

4.5 Remaining Problems

Additional testing, including 2-D offsets, must be performed for Algorithm #1.

5. Potential Change Cueing

The fundamental problem of pixel-level change detection is to utilize multiple image observations of the same scene taken at different times and possible different sensors to identify those pixels in which significant changes have occurred. Changes of military interest include the presence of new man-made features such as roads, bridges and buildings in addition to long-term changes in the boundaries or extent of natural features like forests and rivers.

In an automated change detection system, the primary role of the pixel-level processing is to increase the signal-to-clutter ratio of desired change features to the point where higher level "object" processing can be effective. This requires a high degree of adaptive cancellation of undesired "clutter" due to unchanged or minimally-changed portions of the background scene. If robust background cancellation is achieved, then the potential scene changes can be reliably "cued" on the basis of intensity for use in higher-level analysis.

A key objective of the Phase I effort was to assess how well this objective could be met through the application of multi-image adaptive filtering methods based on statistical decision theory. In the following subsections, we present the basic theory of adaptive change detection, discuss the application of the concept to multi-sensor image data, present the results of change detection experiments conducted under Phase I, and describe some remaining challenges for the Phase II effort.

5.1 Theory of Adaptive Change Detection

A basic theory of detection is formulated below for the case of an arbitrary change signal observed in a multi-image background. A performance measure for pixel-level change enhancement processing is also established.

5.1.1 Modeling

Consider a set of observations $\{x_k\}_{k=1}^N$ from N distinct images. For example, the x_k might represent the samples observed in corresponding co-registered pixels from N different images to be processed for pixel-level changes (although it should be noted that other sets of "observables" can also be used, as noted in Section 5.2 below). If the significant changes in these observations are described as additive signals, then the observation in each image can be modeled by

$$x_k = s_k + n_k, k = 1, \dots, N \quad (5-1)$$

where s_k and n_k represent the change "signal" and the comparatively unchanged background components in the k^{th} image. Using N -vector notation, we can write (1) more compactly as

$$\underline{x} = \underline{s} + \underline{n} \quad (5-2)$$

For modeling purposes, we treat the background \underline{n} statistically by assigning it an arbitrary vector mean $\underline{\mu}$ and an $N \times N$ covariance matrix

$$\underline{K} = E(\underline{n} - \underline{\mu})(\underline{n} - \underline{\mu})^T, \quad (5-3)$$

where "T" denotes the matrix transpose and "E" indicates an ensemble average. This is a commonly used model for natural backgrounds in multi-image remote sensing and signal processing applications. Its utility in describing the pixel variations over local regions in natural scenes has been confirmed.

The change signal \underline{s} is modeled as a deterministic vector. For example, the \underline{s} -vector corresponding to a change feature with additive intensity s_1 present in the first of the N image observations would be denoted by

$$\underline{s} = [s_1, 0, 0, \dots, 0]^T \quad (5-4)$$

The use of a general formulation for the change signal provides flexibility in modeling situations where changes may be present in differing amounts in the multiple images. A typical example is two sets of multi-band imagery collected at two distinct observation times.

Given the above model, it is well-known that the optimum linear filter for the signal \underline{s} is given by

$$\underline{y}(\underline{x}) = \underline{s}^T \underline{K}^{-1}(\underline{x} - \underline{\mu}) \quad (5-5)$$

This multi-dimensional matched filter forms a linear combination of observations which is optimum in the sense that it maximizes the ratio of change signal magnitude to rms background clutter. Note that for the special case where the background statistics are jointly Gaussian, this filter is the maximum-likelihood detector for the signal \underline{s} . For the special case of change detection in a single image with $N-1$ available reference images, the above filter also corresponds to a maximum-likelihood detector derived by [Margalit et al, 85].

In practice, of course, the joint background statistics ($\underline{\mu}$ and \underline{K}) needed to implement the change detection filter are not known in advance, so they must be estimated from observed data by appropriate spatial averaging. This averaging, which forces the filter to adapt to the actual background conditions encountered, is an important aspect of change processing and is discussed in more detail in Section 5.2.

5.1.2 Change Enhancement Performance

The signal enhancement performance of the matched change detector (5-5) is best described in terms of the signal-to-clutter ratio, formally defined by

$$SCR_{out} = \frac{[E\{\underline{y}(\underline{x})\}]^2; \text{ change signal present } (H_1)}{[Var\{\underline{y}(\underline{x})\}; \text{ signal not present } (H_0)]}^{1/2} \quad (5-6)$$

Using (5-3) and (5-5), we have

$$E\mathbf{y}(\underline{\mathbf{x}}) \text{ under } H_1 = \underline{\mathbf{s}}^T \mathbf{K}^{-1} \underline{\mathbf{s}}$$

$$E\mathbf{y}(\underline{\mathbf{x}}) \text{ under } H_0 = 0$$

$$\text{Var}\{\mathbf{y}(\underline{\mathbf{x}})\} \text{ under } H_0 = E\mathbf{y}^2(\underline{\mathbf{x}}) = \underline{\mathbf{s}}^T \mathbf{K}^{-1} E(\underline{\mathbf{x}} - \underline{\mu})(\underline{\mathbf{x}} - \underline{\mu})^T \mathbf{K}^{-1} \underline{\mathbf{s}} = \underline{\mathbf{s}}^T \mathbf{K}^{-1} \underline{\mathbf{s}}$$

The output SCR (5-6) is then

$$\text{SCR}_{\text{out}} = (\underline{\mathbf{s}}^T \mathbf{K}^{-1} \underline{\mathbf{s}})^{1/2} \quad (5-7)$$

For comparative purposes, it is useful to define an input SCR in the k^{th} image prior to multi-image change processing as

$$\text{SCR}_k = |s_k| / \sigma_k \quad (5-8)$$

Where $|s_k|$ and σ_k^2 represent the additive signal magnitude and background clutter (power) in image k , respectively. The multi-image change detection processing gain with respect to bank k is then defined by the ratio

$$G = \text{SCR}_{\text{out}} / \text{SCR}_k. \quad (5-9)$$

To obtain insight into the factors affecting performance, we performed a sample SCR calculation based on equicorrelated background observations in N images. The background covariance matrix for this case is

$$\mathbf{K} = \sigma^2 \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \dots & \dots & \dots & \dots \\ \rho & \rho & \dots & 1 \end{bmatrix} \quad (5-10)$$

where σ^2 is the background power in each image and ρ is the common correlation coefficient between images. The inverse of (5-10) is given by

$$\mathbf{K}^{-1} = \{\sigma^2(1-\rho)\}^{-1} \begin{bmatrix} \alpha & \beta & \dots & \beta \\ \beta & \alpha & \dots & \beta \\ \dots & \dots & \dots & \dots \\ \beta & \beta & \dots & \alpha \end{bmatrix} \quad (5-11)$$

where

$$\alpha = 1 - [\rho / (1 + (N-1)\rho)] \quad (5-12a)$$

$$\beta = -\rho / (1 + (N-1)\rho) \quad (5-12b)$$

The output SCR may be calculated from (5-7) and (5-11)-(5-12) for any specified signal vector $\underline{\mathbf{s}}$. For the change detection application, the signal is often assumed to be present in just one of the images (say, the

first one) and the \underline{s} -vector is given by (5-4). Then we have

$$SCR_{out} = \{s_1 / [\alpha(1-\rho)^{1/2}]\} \{1 - [\rho / (1 + (N-1)\rho)]^{1/2}\} \quad (5-13)$$

The available SCR in the change image alone is

$$SCR_1 = |s_1| / \sigma \quad (5-14)$$

Thus the change detection processing gain for this case is

$$G = SCR_{out} / SCR_1 = \{[(1 + (N-2)\rho) / \{(1-\rho)[1 + (N-1)\rho]\}]^{1/2}\}$$

Fig. 5-1 plots the SCR gain G vs. background correlation ρ for various numbers of images N . The gain is fairly small if ρ is less than 0.9, but very large gains are possible when the background observations are highly correlated ($\rho \rightarrow 1$). It is important to note that most of the available processing gain can be obtained from only 2 image observations, one which contains the change signal and another which serves as the correlated reference for background suppression. The relative change enhancement for the dual-image case is $1/(1-\rho^2)^{1/2}$.

5.2 Application to Multi-Sensor Imagery

The basic principle of pixel-level change detection is relatively straightforward as outlined above. The challenge lies in the application of the techniques to real-world imagery. Key implementation issues which were considered in the Phase I study include:

- 1) The selection of imagery to support change detection processing;
- 2) The selection of the set of "observations" to which change processing is applied;
- 3) The method of averaging used to adapt the change processing to variable background conditions;
- 4) The use of spatial information in the change detection process.

These issues are discussed in turn below.

5.2.1 Selection of Imagery

The analysis in Section 5.1 shows that the most important parameter affecting pixel-level change detection performance is the image-to-image correlation of the background in which the potential changes are to be detected. A higher correlation results in more effective background suppression, which makes small changes easier to discriminate and analyze.

Experience with actual sensor imagery indicates that high values of the pixel correlation coefficient can in fact be obtained from multiple time observations, provided that:

- a) The images are accurately registered to one another;
- b) A close relationship exists between the physical mechanisms that generated the image observations;
- c) The observations are scheduled to minimize decorrelating effects such

as seasonal reflectivity variations, shadowing, solar heating, thermal lag effects, etc.

As a general rule, the observed pixel correlation is strongly related to the phenomenological similarity between various types of images. Thus, for a given scene, optical images tend to correlate well with other optical images, somewhat less well with reflected-IR images, even less with thermal IR images, and quite poorly with radar images. This trend is indicated in Table 5-1 below, which gives the measured correlation coefficient for corresponding 512x512 pixel blocks of pre-processed Raisin City imagery. The reference band TM1 is located in the blue-green optical region from 0.45-0.52 microns.

Waveband Pair	Comments	Correlation Coefficient
TM1 x TM3	optical x optical (0.63-0.69 μ m)	0.98
TM1 x TM5	optical x near-IR (1.55-1.75 μ m)	0.89
TM1 x TM6	optical x thermal-IR (10.40-12.50 μ m)	0.36
TM1 x SAR	optical x L-band aircraft radar	0.12

Table 5-1: Correlation Coefficients for Raisin City Images

On this basis, one would predict that good pixel-level change enhancement could be obtained by pairing a TM1 image with a TM3 image, but that little enhancement would result from processing corresponding pixels in the TM1 and SAR images. Change detection experiments discussed in Section 5.3 generally confirm this expectation.

Even when multiple images from the same spectral band are available for change detection, diurnal and seasonal variations in the imaged intensity values can decorrelate the observed data. Although certain seasonal variations may constitute changes of military interest, the presence of too much image-to-image variation can make it very difficult to achieve effective image-to-image registration and background suppression. At thermal infrared wavelengths, for example, images taken at night often differ considerably from daytime images due to the lack of a solar reflected component and the presence of temperature differentials resulting from the variable thermal lags of different materials. Seasonal variations are induced by changes in the moisture content of the terrain or vegetation; i.e. damp soil normally has a lower apparent temperature than dry soil due to the effect of evaporative cooling.

Certain components of long-wave IR images of natural scenes can be fairly well correlated, even when taken years apart. For example, Table 5-2 show the measured correlation coefficients for various image pairs from the Death Valley scenes collected in 1983 and 1988. Bands 1, 3 and 5 are the

original radiance images after global registration. The "PC" images are the three principle spectral components generated from Bands 1, 3 and 5 and shown originally in Section 2.4

Image Pair	Correlation
TIMS Band 1	0.31
TIMS Band 2	0.32
TIMS Band 3	0.40
PC Image 1	0.27
PC Image 2	0.89
PC Image 3	0.59

Table 2: Pixel Correlation of TIMS Death Valley Data Taken in 1983 and 1988

Note that relatively high pixel correlations are observed in the second and third principal components of the scene.

For change detection within a single waveband, the use of multi-band long-wave thermal imagery in the 8-12 μ m region appears to have considerable potential. Although in general terrain features both emit and reflect IR radiation, the dominant phenomenon in the long-wave region is greybody emission as a function of object temperature and emissivity. Since undisturbed natural backgrounds exhibit relatively little long-term change in emissivity, the major parameter that determines the apparent radiance is the temperature. High correlations between thermal images taken at different times can be expected if the observations are carefully scheduled to minimize differential temperature variations due to diurnal and seasonal effects.

5.2.2 Selection of Observations

The most natural observations to combine in a change detection operation are the corresponding pixel values in the registered images. In this case, the pixel correlation coefficient determines the relative amount of change enhancement that can be achieved.

In principle, however, one can utilize any set of observables that can be derived from the image pixel data. An interesting alternative which was examined in the Phase I study is the use of image spatial frequency components, rather than image pixels, as the input "data" to the change detection process. This particular choice was motivated by the empirical observation that much of the correspondence between displayed images from different spectral bands resides in the edges and other high frequency artifacts. If the frequency components corresponding to these features can be isolated by means of Fourier analysis, their higher correlations could be exploited to achieve more effective clutter cancellation.

Fig. 5-2 shows a portion of the registered TM1 and Aircraft SAR images from the Raisin City site. As noted above, the average pixel correlation coefficient for these images is only 0.12, implying that they are practically uncorrelated at the pixel level. This is not surprising in view of the different sensors used and the decorrelating effect of the SAR image speckle. However, the correspondence between some of the boundaries in these scenes is evident.

To examine the sensitivity of the correlation coefficient to spatial frequency for this image pair, the following procedure was used. First, the two full TM1 and SAR images were sub-divided into 16 128x128 pixel sub-blocks, a 2-D DFT of each block was computed, and the complex components of the corresponding blocks from each image were conjugate multiplied together (i.e., correlated). The like-indexed correlations in all blocks were then averaged together and normalized to obtain an ensemble average correlation coefficient estimate as a function of 2-D spatial frequency. The magnitude of the resulting correlation array is plotted in grey-scale format vs. normalized 2-D spatial frequency in Fig. 5-3(a), where black indicates a moderate positive correlation (0.6) and white denotes zero correlation. The histogram of spatial frequency correlations, shown in Fig. 5-3(b), has a mean value of 0.23. Although many of the spatial frequencies are almost as poorly correlated as the pixels, there are some notable exceptions. For example, most of the components in the vicinity of zero horizontal ("X") frequency (oriented vertically in the plot) have a moderately high correlation (0.6 or greater). These components are dominated by edges with a near horizontal orientation in the scene.

Change detection can be implemented with spatial frequency domain in essentially the same way as for image pixels, with the necessary modifications for complex-valued data. Given a complex vector \underline{x} representing a set of corresponding spatial frequency components from N co-registered image blocks, we compute the linear filter.

$$y(\underline{x}) = \underline{s}^H \underline{K}^{-1} (\underline{x} - \underline{\mu}) \quad (5-16)$$

where \underline{s} is the change signal, ' H ' denotes the conjugate transpose, and $\underline{\mu}$ and \underline{K} now represent the background mean and covariance at a single spatial frequency of interest. This same process can be performed for all available frequency components in each block, after which the results can be filtered and transformed back to the image domain for interpretation.

5.2.3 Adaptation to the Background

Although the matched change detection filter described in Section 5.1 is an optimum linear discriminant, its immediate application is limited by the apparent need for prior descriptions of the scene background statistics (i.e., mean and covariance). A standard way to remove this dependence is to estimate the ensemble statistics by averaging over a sufficiently large number of observations. If the changes to be discriminated are small in magnitude or occupy only a small fraction of the observations chosen, then their contribution to these averages will be

small and can be ignored for all practical purposes.

For pixel change processing, the simplest approach to adaptivity is to perform background averaging and subsequent filtering over the entire scene. A potential problem with this method is that the background statistics can vary considerably with changes in local surface content, leading to local change filter mismatch. The use of small spatial windows alleviates the nonstationarity problem at the expense of a reduced sensitivity to extended change features. For dual-band change detection, it can be shown that a covariance estimate based on at least 49 vector samples (i.e., a 7x7 window) provided detection sensitivity to within 1dB of the perfectly matched filter. However, the smaller the window, the more likely it is that larger sized changes will dominate the averages taken in particular windows, resulting in self-suppression of these features during the matched filtering operation. If change features with a very wide range of sizes and shapes must be discriminated, then the use of a pyramid of processing window sizes may be appropriate. A more computationally-intensive option is to perform brightness and texture-based scene segmentation prior to change processing, to identify those regions which appear to have nearly homogeneous background statistics.

For frequency component change processing, the assumption is that the like components taken from the different blocks of an image are drawn from a stationary ensemble, but that the statistics of different components can vary arbitrarily. The averaging is therefore performed component by component across the DFT blocks in the image. For a fixed size image, a fundamental tradeoff exists between spatial frequency resolution (which improves with a larger block size) and the number of independent blocks available in the image (which decreases with the use of larger blocks).

5.2.4 Use of Spatial Information

A purely pixel-level change detection process utilizes the information available in corresponding pixels of two or more images, leaving it up to the interpreter to make further inferences based on feature size, shape and orientation. However, it would be possible to optimally combine information from multiple pixels for automated change detection based on a feature "shape" hypothesis. A general formulation of this approach, utilizes a pixel-level change filter followed by a matched spatial filter. Although the proper use of shape information could lead to higher signal-to-clutter ratios for change cueing, several potential problems arise. The first results from the wide range of potential feature sizes and shapes, which would require the use of a large bank of change filters matched to various objects. Secondly, the preferential spatial filtering performed by this approach would alter the appearance of objects in the scene, making higher level contextual interpretation more difficult.

Since frequency component change processing is carried out in the Fourier transform domain, it is easy to modify the spatial frequency content of the change image, if desired, through FIR filtering. The frequency-domain equivalent of the spatial enhancement approach described above would involve the application of a frequency-dependent weighting which is proportional to the spectrum of the feature of interest. However, with no

prior spatial information about the features, it appears that a better approach is to simply normalize all frequency components so that they contribute the same relative power to the final change image. Since the power in each spectral change component is equivalent to its variance, such normalization can be achieved by modifying (5-16) to implement

$$y(x) = S^H K^{-1} (x - \mu) \quad (5-16)$$

at each frequency in the band. This spectral "whitening" appears to be quite effective for change detection, as seen in the examples below.

5.3 Change Detection Results

The change detection concepts discussed above were qualitatively tested by processing two different sets of imagery. Preprocessed Raisin City TM and SAR images with synthetically injected changes were used to examine the performance of variations of change enhancement processing. Although the Raisin City multi-band data set does not contain long-term changes, it is still useful for illustrating the effectiveness of change processing applied to imagery taken from different bands. Sensitivity to actual long-term changes was examined by processing TIMS Death Valley long-wave infrared data collected in 1983 and 1988.

5.3.1 Raisin City Data

The Raisin City multi-band data set consisted of 7 Landsat TM bands and 3 SAR images co-registered to one another. The four images shown in Table 5-3 were selected for change processing experiments.

Band Designation	Comments
TM1	optical band (0.45-0.52 μ m)
TM3	optical band (0.63-0.69 μ m)
TM5	near-infrared band (1.55-1.75 μ m)
SAR	JPL Aircraft SAR (L-band)

Table 5-3: Raisin City Images Used In Change Detection Experiments

The TM1 optical band was selected as the "change image". For reference purposes, its global histogram was normalized to zero mean and unit variance. Change features of varying size and brightness were then synthetically introduced into the normalized TM1 image using the four-step procedure illustrated in Fig. 5-4. The first step was to inject the following change signatures shown in Fig. 5-4(a) by direct pixel intensity replacement:

- a) Small Structures. A set of three 2x2 pixel features with negative local contrast and average replacement amplitude -1;

- b) Road. A linear feature of width 2 pixels and average pixel replacement amplitude 1, corresponding to the mean intensity of a diagonally-oriented road in the scene;
- c) Large Structure. A rectangular 8x4 pixel feature with positive local contrast and average replacement amplitude 3;
- d) Field. A rectangular feature of size 30x35 pixels with average pixel intensity 2, comparable to that of several other brightly-reflecting fields in the scene.

The actual pixel amplitudes of the features were randomly dithered about their average values to simulate the local rms intensity variation observed in the TM1 scene.

The remaining three steps shown in Fig. 5-4 were used to model the effect of finite-resolution TM optics on the injected feature responses. The additive change image shown in Fig. 5-4(b) was first extracted by subtracting the frame shown in Fig. 5-4(a) from the TM1 frame prior to injection. Next, the additive change signals were convolved with a 2-D Gaussian function having a 1-sigma blur circle of 2 pixels to produce the frame in Fig. 5-4(c). Finally, the blurred signals were added back into the original TM1 image to generate a final scene containing injected features with more realistic spatial frequency content, as shown in Fig. 5-4(d).

Fig. 5-5 shows the full 512x512 TM1 frame with injected changes, referred to below as TM1C. A number of change detection experiments were conducted using the TM1C image paired with other Raisin City images in Table 5-3 to determine how well the change features could be enhanced for reliable cueing on the basis of brightness; typical results are presented below. It should be emphasized that all of the change images shown are normalized and presented on a common grey-scale, so that direct visual comparisons among them can be made.

The TM1C-TM3 pair is used to illustrate several variations of the change processing. The first experiment consisted of applying the change detection filter to each image pixel, using global averages of the background mean and covariance computed over the entire 512x512 pixel scene. The resulting output frame is shown in Fig. 5-6. The synthetic change features are obviously enhanced relative to Fig. 5-5 and their shapes are well preserved. Several other portions of the scene are also enhanced by the process. However, it is difficult to determine whether other bright regions correspond to significant changes, or simply result from poor background suppression due to local mismatch in a change filter based on global scene averages. The "patchiness" of the change image in Fig. 5-5 is a direct result of the nonstationary background in the Raisin City scene.

An obvious way to improve the ability of the change filter to adapt to local background conditions is to utilize smaller averaging windows. Fig. 5-7 shows a typical result for pixel-level change processing based on an averaging block of size 16x16 pixels. In each 16x16 block, the local background mean and covariance were used to compute the change filter to be applied to the 256 pixels of that block, after which the output was

variance-normalized for presentation on a common grey-scale. The uniformity of the suppressed background in Fig. 5-7 is an improvement over Fig. 5-6, but the smaller averaging window clearly results in significant self-suppression of the change features; only the road is easily discernable.

The results of a frequency-component change detection approach are shown in Fig. 5-8. To generate this image, the bands TM1C and TM3 were each subdivided into 16 blocks of size 128x128 pixels which were transformed to the spatial frequency domain by 2-D DFTs. The adaptive change filter (5-17) was then applied to each frequency component of every block, based on background statistics averaged over the like components in all blocks. Finally, the processed components of each block were transformed back to the image domain via 2-D inverse DFTs and mosaiced to form the output change image shown in Fig. 5-8. Note that all of the synthetic changes stand out very well against the suppressed background. Evidently, the frequency domain technique can provide superior local background suppression without excessive attenuation of larger change features. This is a very valuable attribute in situations where a wide range of feature sizes and shapes are of interest.

Other experiments were performed to confirm the original expectation that TM3 is the best reference band for change detection in TM1. For example, Fig. 5-9 and 5-10 show the results of frequency component change processing applied to the image pairs TM1C-TM5 and TM1C-SAR, respectively. The results obtained with the infrared band (TM5) in Fig. 5-9 are still rather good, but the overall background suppression performance is not as high as in Fig. 5-8 due to the lower correlation between the optical and infrared bands. The results for the SAR image in Fig. 5-10 are poorer still, with increased edge leakage and almost no enhancement of the large "field" feature.

6. Preliminary Analysis of Effects of Variations in Imaging Scenarios

One concern when creating algorithms for registration and change detection is the robustness of such procedures over the varieties of imaging scenarios. This section contains some observations on imaging variabilities for SAR in 6.1 and optical sensors in 6.2.

6.1 Variable SAR Scenario Effects

Radar returns from vegetation canopies consist of scattering from the vegetation volume, the soil, and the soil-vegetation interaction. These terms depend on the foliar and woody biomass and the soil state. Near vertical, the soil term becomes dominant. For higher incidence angles, the volume term dominates for lossy canopies.

Horizontal (H) polarization couples weakly with vertical stalks, whereas the opposite is true for vertical (V) polarization. Therefore, as a general rule, the H polarization reveals more about soils and the V polarization reveals more about canopies. As always for a given set of radar parameters, the primary determinant of attenuation in vegetation is water/unit volume.

Co-polarized returns are usually stronger than cross-polarized returns. The latter are caused by four basic mechanisms [Fung, Ulaby,83]:

- o Polarization dependence of Fresnel coefficients for quasi-specular reflection;
- o Multiple surface scattering induced by surface roughness;
- o Multiple volume scattering due to inhomogeneities within the skin depth;
- o Physical or geometric target anisotropy.

One measure of the degree of inhomogeneity is the ratio $(\sigma_e/\lambda)^2$, where σ_e is the standard deviation of the dielectric constant in the boundary layer of surface scattering or of the spatial discontinuities in volume scattering. Therefore, depolarization increases with increasing frequency.

Also cross-polarization (HV and VH) returns have a weaker angular dependence than like polarization (HH and VV) returns, especially near the vertical. Therefore, the cross-polarization ratio σ_{HV}/σ_{HH} increases with increasing angle and is larger for crop canopies than for bare soil. Therefore, cross polarization is generally more suitable for slope studies.

6.2 Variable Optical Scenario Effects

The quantity that is directly measured in well-calibrated electro-optical imagery is the in-band pixel apparent radiance. In general, this "sensed"

radiance exhibits a complex dependence on two major classes of parameters:

- 1) Physical properties of the terrain features being imaged;
- 2) Sensor, environmental and encounter parameters specific to the imaging scenario.

In the optical and near-IR region, the primary physical attribute of a material is its reflectance spectrum. In-band reflectance is conventionally measured from remotely-sensed multi-band imagery by postulating the existence of a uniform illumination function across the scene and normalizing it out of the data. A large body of experience with actual multi-spectral imagery (e.g., Landsat) has shown that measured reflectance spectra of terrain regions are subject to significant long-term changes. For example, the reflectance of healthy ground vegetation peaks strongly in the near-infrared bands due to relatively high moisture content, while for stressed vegetation this effect is far less pronounced. At optical wavelengths, the presence of a surface moisture layer on opaque materials tends to reduce the measured reflectivity.

Sensed radiance in the thermal infrared is a function of material properties (apparent temperature and emissivity) as well as scenario-dependent parameters (atmospheric path transmission, path radiance, solar reflection, etc.). Image preprocessing can be used to factor out the effects of certain scenario-dependent parameters in order to obtain physical measurements (i.e., temperature). Standard tools such as the LOWTRAN7 computer code, a comprehensive atmospheric model, are available for this purpose. Change analysis is complicated not by bulk variations in temperature across a scene during the normal course of a day, but rather by differential temperature changes among materials with variable thermal inertias.

The encounter geometry also influences electro-optical imagery, although to a lesser extent than microwave radar images. Most natural surfaces are diffuse reflectors and emitters of optical radiation, and can be considered Lambertian to a first approximation. The reflectivity and emissivity of such materials do not show a strong angular dependence. On the other hand, features like man-made metal surfaces and smooth bodies of water can appear nearly specular at optical and infrared wavelengths; their signatures vary considerably with the imaging aspect and the solar angle. Overhead electro-optical imagery used for terrain surface characterization is often collected early or late in the day to minimize specular reflection effects.

Environmentally-induced variations in electro-optical imagery result from the spectral content of the illumination and local atmospheric path conditions (molecular constituents, temperature, moisture content, etc.). For example, direct solar illumination has different spectral properties than indirect illuminations due to blackbody emission from clouds. Atmospheric conditions can generally be predicted as a function of local meteorological conditions, time of year and earth latitude and longitude.

The primary sensor parameter affecting observed electro-optical imagery is the waveband of operation, since it determines the basic phenomena being

measured (i.e., reflection, emission, or both). Secondary sensor-induced variations result from design and sensitivity differences among instruments, such as scanning vs. staring arrays, detector sensitivity and spacing, system noise levels, and so on. An important sensor-specific effect that frequently impedes electro-optical image processing is pattern noise resulting from detector responsivity variation across the focal plane array, as well as gain and offset imbalances in the A/D converters used to digitize the focal plane readout.

7. Conclusions

An overview of the results obtained during the Phase I effort and remaining problems appears in section 7.1. Some recommendations for the Phase II effort are discussed in section 7.2.

7.1 Summary of Results and Remaining Problems

The image-image registration efforts were initially encouraging. The two types of rough registration methods were area-based and contour-based. These algorithms operated under the assumption of little terrain distortion to the imagery. An contour-based algorithm for dealing with the problem of considerable terrain relief is outlined. Fuller development and testing of this procedure is a Phase II issue.

A two-stage method of combining the area-based and contour-based algorithms showed some improvement over the performance of each separately. In particular, small residual amounts of rotation on the order of a degree were removed by this procedure.

Two sub-pixel registration estimation algorithms are also presented, with extensive testing results so far available for one of them. The further investigation of the other algorithm will be undertaken in Phase II.

The following statistics summarize registration accuracies achieved in testing:

- o area-based
 - K-L algorithm: ± 1 pixel in 50% of cases,
 - MNF algorithm: ± 1 pixel in 75% of cases,
- o contour-based: ≤ 2 pixels in all cases,
- o combined area-contour: ± 1 pixel (using MNF area method)
- o sub-pixel registration:
 - Algorithm #1: $\sim 5 \times 10^{-4}$ pixel (using lower frequencies, and without spectral leakage filtering)
 - Algorithm #2: $< 10^{-3}$ pixel (best results using Hanning filter for filtering spectral leakage))

The sub-pixel results were on simulated data.

The other main technical area investigated was change cueing on the pixel level, given a registered image pair. Both actual changes and simulated changes were examined. The results of testing revealed excellent cueing even for smaller targets as long as the local image-image correlation of the background was high, with the performance degrading as this correlation decreases.

The remaining problems/concerns are:

- o Obtaining a wider variety of multi-sensor image pairs for different types of scenes and sensor parameters,
- o More extensive testing and characterization of performance of present methods.

7.2 Recommendations for Phase II

Phase II should continue the development and testing of automated methods for:

- o Area-based rough registration.
- o Contour-based registration,
 - present method for imagery with little terrain,
 - new method for terrain-distorted imagery.
- o Sub-pixel registration method #1.
- o Change cueing,
 - usage of physically derived quantities in EO and IR imagery,
 - frequency space analysis for determining background correlation,
 - target model, object-based analysis.

From a signal processing standpoint, this performance dependence on the background correlation levels is unavoidable for pixel level processing. However, considerably more progress can be achieved using processing methods which effectively increase these background correlation levels by restricting attention to selective frequency regions.

Progress beyond what can be achieved using such enhanced pixel-level processing would probably require higher-level procedures on the object-level. Such methods will be required to make hypotheses on the existence of objects based on pattern analysis, as opposed to simply thresholding based on local statistics. The investigation of such object-based target cueing methods will require the use of target models.

For optical and IR imagery, it is clear that the performance of algorithms for sub-pixel registration and change detection will depend heavily on the degree of correlation between images. Images based on a direct observable such as pixel apparent radiance can be significantly decorrelated over the long-term changes in the imaging geometry, season, time of day, amount of solar reflection, etc. However, it appears that the effects of certain variations can be reduced in many cases by converting the imagery to a more fundamental physical quantity prior to change detection.

In the thermal infrared, for example, the sensed radiance is a function of actual material properties such as apparent temperature, and emissivity, as well as scenario-dependent parameters such as atmospheric path transmission, path radiance, solar reflected component, etc. The ultimate goal of the pre-processing is to factor out the effects of the scenario-dependent parameters to obtain physical measurements, such as temperature, that may exhibit a greater long-term correspondence between well-scheduled

image observations. Available tools such as LOWTRAN7 can be used for this purpose during Phase II.

Despite the encouraging results using the automated methods, it is strongly recommended that the Phase II workstation contain capabilities for interactive as well as automated modes of operation. Such a dual capability allows the use of a human operator to examine and assess the results generated by automated registration procedures, make corrections if needed, and to provide initial offsets for difficult or ambiguous cases.

In particular, VEXCEL has developed, on another effort, software for an electronic light table which has all of the capabilities of a hardcopy light table. This electronic light table (LT) is sensor independent. Together with a sensor model, LT provides all of the image manipulation, visual enhancement, and editing facilities for image registration.

The inclusion of VEXCEL's LT software package into the Phase II prototype will provide the needed interactive image registration capabilities, which can then be interface with the library of automated techniques.

For change detection, the use of automated change cueing requires only selective attention by the operator, but uses his/her superior judgement for evaluating cues as legitimate targets. Therefore, LT's visual image enhancement tools can be used for examining such cued potential targets

8. References

- o Adair, M., Guidon, B., "Statistical edge detection operators for linear feature extraction in SAR images," Can. J. Rem. Sens., vol-16, #2, July 1990.
- o Aho, A. Design and Analysis of Computer Algorithms, Addison Wesley, 1974.
- o Ayache, N. et al., "Efficient Registration of Stereo Images by Matching Graph Descriptions of Edge Segments," Int. J. Comp. Vision, Vol. 1, #2, 1987.
- o Ayache, N. et al., "A New Method for the Recognition and Positioning of 2-D Objects," ICPR, Montreal Canada, 1984.
- o Ballard, D. et al., Computer Vision, Prentice Hall, 1982.
- o Barnea, D.I., Silverman, H., "A Class of Algorithms for Fast Image Registration," IEEE Trans. Comput., Feb. 1972.
- o Barrows T., et al., "Parametric Correspondence and Chamfer Matching: Two New Techniques for Image Matching," Proc. DARPA Image Understanding Workshop, May 1978.
- o Blom, R.G. et al., "Radar satterometry of sand dunes and lava flows," Proc. IGARSS'82, Munich, June 1982.
- o Blom, R.G. et al., "Results from OSTA-1: Sir-A images of volcanic fields and sand dunes," Proc. IGARSS'82, Munich, June 1982.
- o Blom, R.G. et al, "What are best radar wavelengths, incidence angles and polarizations for geological applications? A statistical approach," Proc. IGARSS'85, New York, 1985.
- o Blom, R.G. et al, "What are best radar wavelengths, incidence angles, and polarizations for discrimination among lava flows and sedimentary rocks? A statistical approach," IEEE Trans. Geosci. Rem. Sens., vol-25, 1987.
- o Blom, R.G., Elachi, C., "Multifrequency and multipolarization radar scatterometry of sand dunes and comparison with spaceborne and airborne radar images," JGR, vol-92, 1987.
- o Blom R.G., Elachi, C., "Spaceborne and airborne imaging radar observations of sand dunes," JGR, vol-86, 1981.
- o Blom, R.G. et al., "Detection of subsurface features in Seasat radar images of Means Valley, California," Geology, vol-12, 1984.
- o Blom, R.G. et al., "On the relationship between age of lava flows and radar backscatter," Proc. IGARSS'86, Paris, 1986.

- o Blom, R.G., "Effects of variation in look angle and wavelength in radar images of volcanic and aeolian terrains, or Now you see it, now you don't," Int. J. Rem. Sens., vol-9, #5, 1988.
- o Bolles, R. et al., "Recognizing and Locating Partially Visible Objects: The Forms Feature Method," Intl. J. Robot. Res., Vol. 1, #3, 1982.
- o Bruckstein, A., Netravali, A., "On Minimal Energy Trajectories", Computer Vision, Graphics, and Image Processing, vol 49, 1990.
- o Castleman, K., Digital Image Processing, Prentice-Hall, 1979.
- o Champeney, D., Fourier Transforms and their Physical Applications, Acad. Press, 1973.
- o Curlander, J., Kober, W., "Smart Real-Time Control Generator," Final Report, Option 1&3, Phase II SBIR, Contract # DACA72-87-C-001, U.S. Army Engineer Topographic Laboratories, Ft. Belvoir, VA., 1989.
- o De Castro, E. et al., "Compensation of Random Eye Motion in Television Ophthalmoscopy: Preliminary Results," IEEE Trans. on Med. Image., MI-6, #1, Mar 1987.
- o Frei, W., "Image Enhancement by Histogram Hyperbolization," J. Comp. Graph. Image Proc., 1977.
- o Frost, V. et al., "A Model for Radar Images and its Application to Adaptive Digital Filtering of Multiplicative Noise," IEEE Trans. Patt. Anal. Mach. Intell., vol-4, #2, Mar. 1982.
- o Fullerton, K.J., et al., "Opposite-Side SAR Image Processing for Stereo Viewing," Photog. Eng. Rem. Sens., vol-52, #9, 1986.
- o Fulks, W., Advanced Calculus, Wiley, 1966.
- o Fung, A., Ulaby, F., "Matter-energy interaction in the microwave region," in: Manual of Remote Sensing, Amer. Soc. Photog., 1983.
- o Gotashby, A., "Registration of Images with Geometric Distortions," IEEE Trans. Geo. Rem. Sens., vol. 26, #1, Jan. 1988.
- o Green, A., et al., "A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal," IEEE Geo. Rem. Sens., vol-26, 1988.
- o Hu, M., "Visual Pattern Recognition by Moment Invariants," IRE Trans Inform Thy., vol-8, Feb. 1962.
- o Huang, T. ed. Image Sequence Analysis, Springer, 1981.

- o Kalvin, A., et al., "Two-Dimensional Model-Based Boundary Matching using Footprints," Intl. J. Rob. Res., Vol. 5, #4, Winter 1986.
- o Kober, W., et al., "A Smart Mapping, Charting, and Geodesy Control Generator," Final Report, Base Contract, Phase II SBIR, Contract # DACA72-87-C-001, U.S. Army Engineer Topographic Laboratories, Ft. Belvoir, VA., 1988.
- o Kuglin, C.D. et al., "The Phase Correlation Image Alignment Method," Proc. IEEE 1975 Conf. on Cyber Soc., Sept. 1975.
- o Kuglin, C.D. et al., "Map Matching Techniques for Terminal Guidance Using Fourier Phase Information," Proc. SPIE: Digital Processing of Aerial Images, Vol. 186, 1979.
- o Lahart, M.J., "Optical Area Correlation with Magnification and Rotation," J. Opt. Soc. Amer., Vol. 6D, May 1970.
- o Lancaster, P., Salkauskas, K., Curve and Surface Fittings: An Introduction, Acad. Press, 1986.
- o Leberl, F. et al., "Multiple Incidence Angle SIR-B Experiment over Argentina: Stereo Radargrammetric Analysis," IEEE Trans. Geo. Rem. Sens., GE-24, #4, 1986.
- o Lee, J.S., "Digital Image Smoothing and the Sigma Filter," Comp. Vis, Image Proc., vol-24, 1983.
- o Lucke, R., "A Local Interpolator Derived from the DFT," Naval Research Laboratory Report (undated).
- o Lukert, D., Blanchard, A., "Azimuth depolarization ambiguities in Synthetic Aperture Radar," Int. J. Rem. Sens., vol-9, #3, 1988.
- o Maitra, S., "Moment Invariants," Proc. IEEE, vol-67, April 1979.
- o Margalit, A., Reed, I.S., Gagliardi, R.M., "Adaptive Optical Target Detection Using Correlated Images," IEEE Trans. Aerosp. Elect. Sys., vol-AES-23, #1, May 1985.
- o Medioni, G. et al., "Matching Images using Linear Features," IEEE Trans. Patt. Anal. Mach. Intell., PAMI-6, #6, Nov. 1984.
- o McConnell, R., "Matching of Dissimilar Radar Images Using Marr-Hildreth Zero Crossings," Proc. Amer. Soc. photog. Rem. Sens., Balt. MD, April 1987.
- o Nagao, M., "A Nonlinear Edge Preserving Smoothing Algorithm," Comp. Graph. Image Proc., vol-9, 1979.
- o Nevatia, R., Machine Perception, Prentice Hall, 1982.
- o Papoulis, A., Signal Analysis, McGraw Hill, 1977.

- o Pearson, J.J. et al., "Video Rate Image Correlation Processor," Proc SPIE: Applications of Digital Image Processing, Vol. 119, 1977.
- o Pratt, W. "Correlation Techniques of image Processing," IEEE Trans. Aerosp. Elect. Sys., May 1974.
- o Pratt, W., Digital Image Processing, Wiley, 1978.
- o Price, K., "Symbolic matching of images and scene models," Proc. IEEE Workshop Comput. Vision, Aug. 1982.
- o Ramapriyan, H.K. et al., "Automated Matching of Pairs of SIR-B Images for Elevation Mapping," IEEE Trans. Geo. Rem. Sens., vol-24, #4, July 1986.
- o Roberts, R.A., Mullis, C.T., Digital Signal Processing, Addison Wesley, 1987.
- o Rosenfeld, A., et al., Digital Picture Processing, Acad. Press, 1982.
- o Rosenfeld, A., ed. Multiresolution Image Processing and Analysis, Springer Verlag, 1984.
- o Sankoff, et al., Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison, Addison Wesley, 1983.
- o Schaefer, R.W., Rabiner, L.R., "A Digital Signal Processing Approach to Interpolation," Proc. IEEE, vol-61, #6, 1973.
- o Schaum, A., McHugh, M., "New Analytic Methods of Image Registration," presented at the 36th National Infrared Information Symposium, NBS, Boulder, CO, May 1988.
- o Schwarz, J. et al., "Identification of Partially Obscured Objects in Two and Three Dimensions by Matching Noisy Characteristic Curves," Robotics Report #46, New York Univ., 1985.
- o Shapiro, L. et al., "Structural Descriptions and Inexact Matching," IEEE Trans. Patt. Anal. Mach. Intell., PAMI-3, Sept. 1981.
- o Stocker, A.D., Clayton, N.W., "A Discrete Phase Correlation Algorithm for Frame Registration," Space Computer Corporation Report SCC-R-164-4, July 1989.
- o Stocker, A.D., "Phase Correlation and Optimum Displacement Estimation," Space Computer Corporation Report SCC-R-181-3, October 1990.
- o Svedlow, M., et al., "Image Registration: Similarity Measure and Preprocessing Method Comparisons," IEEE Trans. Aerosp. Elect. Sys., AES-14, #1, Jan. 1978.

- o Turney, J. et al., "Recognizing Partially Hidden Objects," IEEE Intl. Conf. Robot. and Autom., 1985.
- o Wolberg, G., "Digital Image Warping," IEEE Computer Society Press, Los Alamitos, CA, 1990.
- o Wong, R., "Sensor Transformations," IEEE Sys. Man Cyb., vol-7, 1977.
- o Wong, R., "Intensity Signal Processing of Images for Optical to Radar Scene Matching," IEEE ASSP, vol-28, 1980.
- o Zebker, H. et al., "Imaging Radar Polarimetry From Wave Synthesis," J. Geophys. Res., vol-92, #B1, Jan. 1987.

9. Appendix

9.1 Additonal Imagery

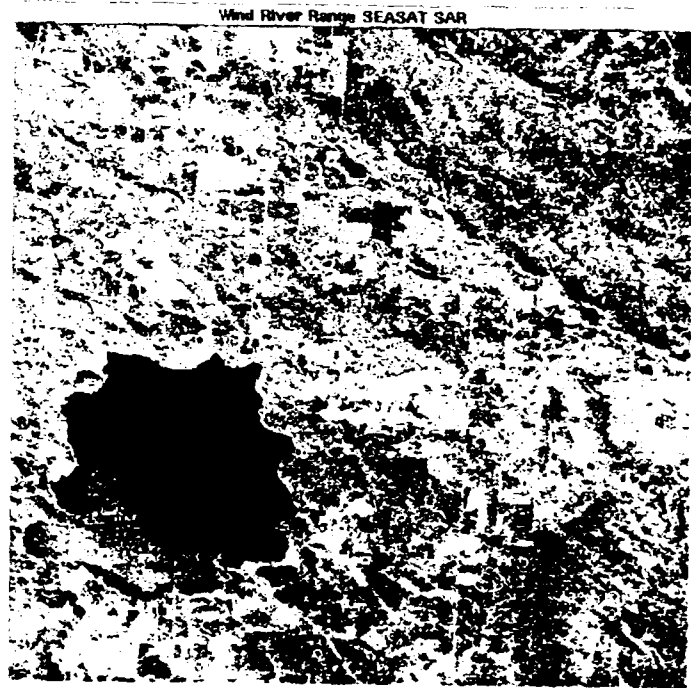
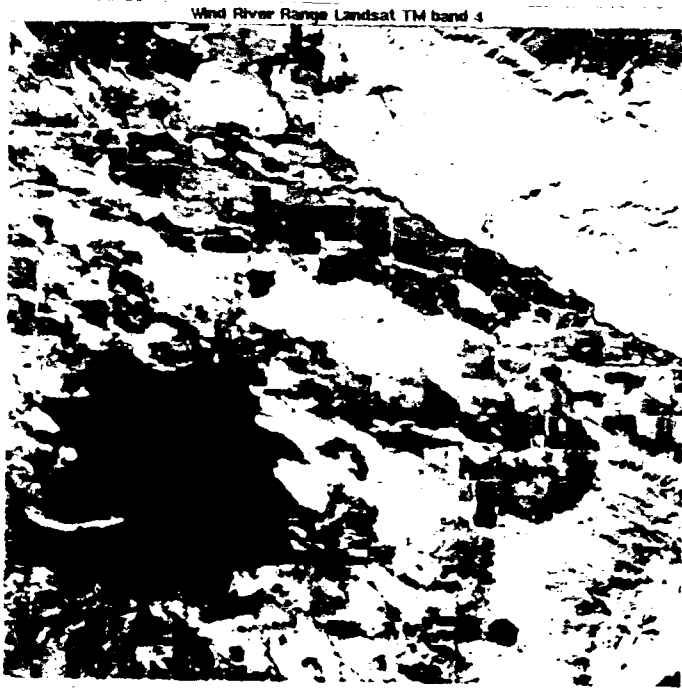


Fig. 9-1(a): Wind River Basin, WY



Fig. 9-1(b): Altamaha River GA Region

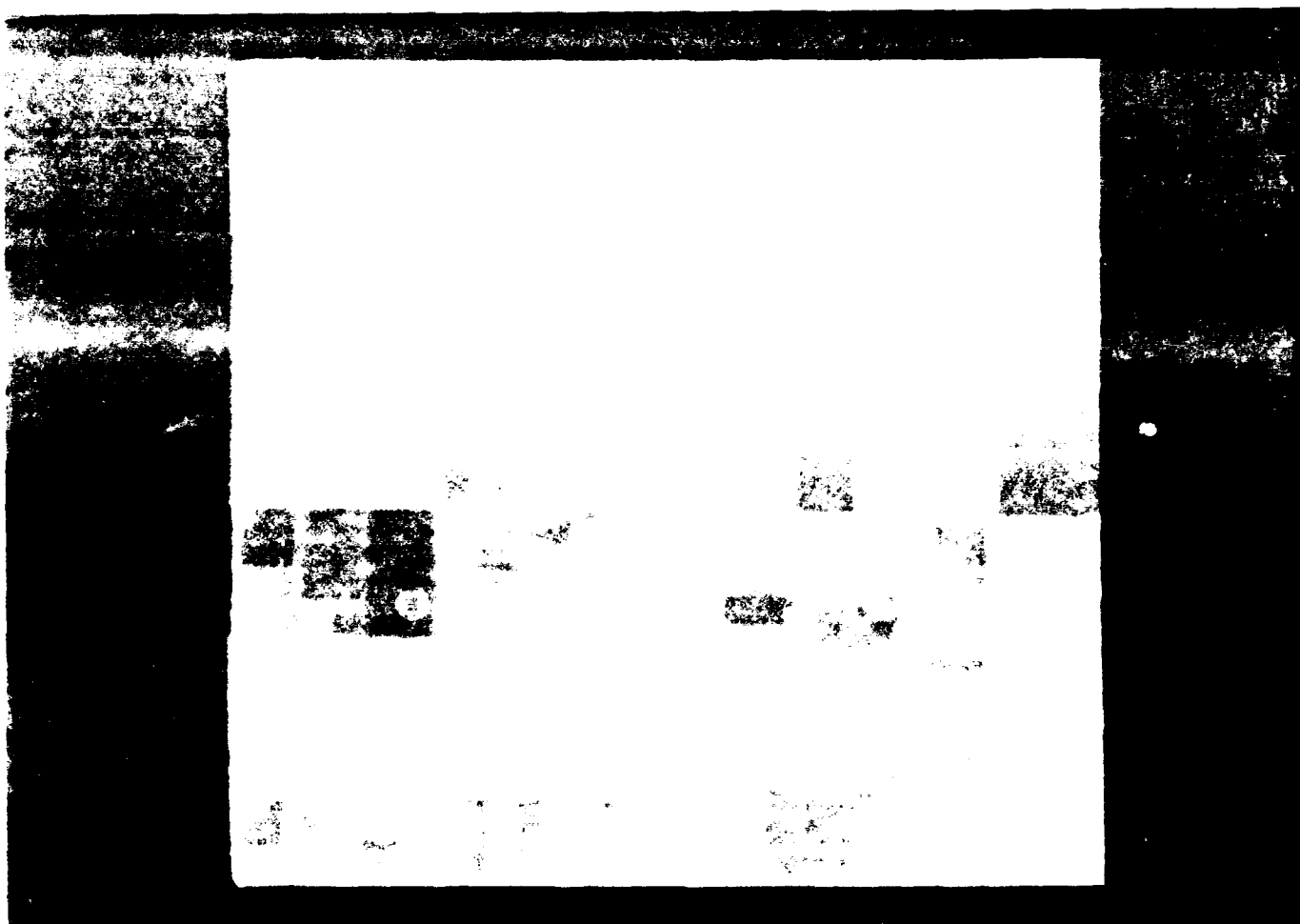


Fig. 1. View of the road from the camp, Raisin City, CA

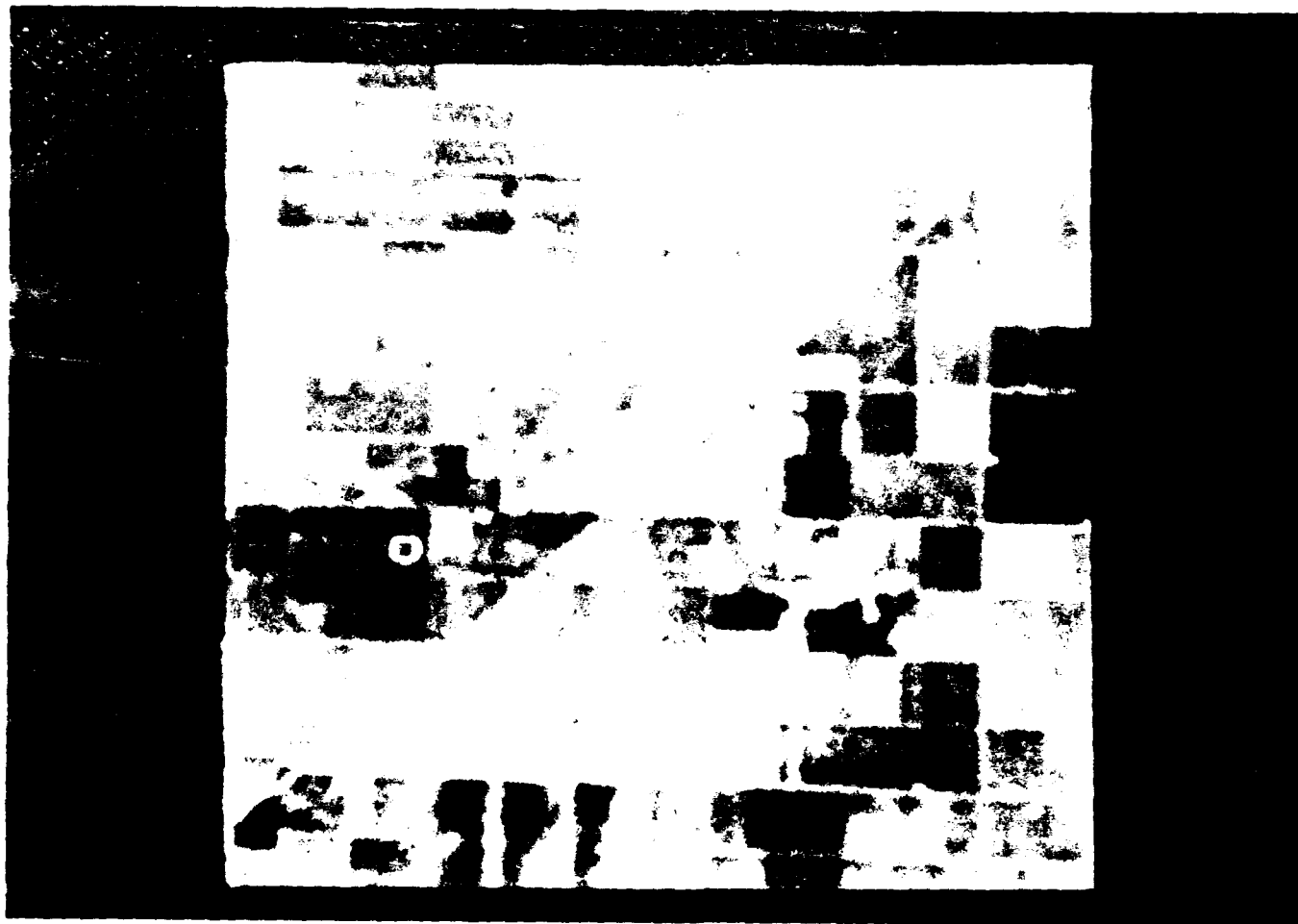


Fig. 9-3: Landsat TM Band 2 (.52 μ m-.60 μ m), Raisin City, CA

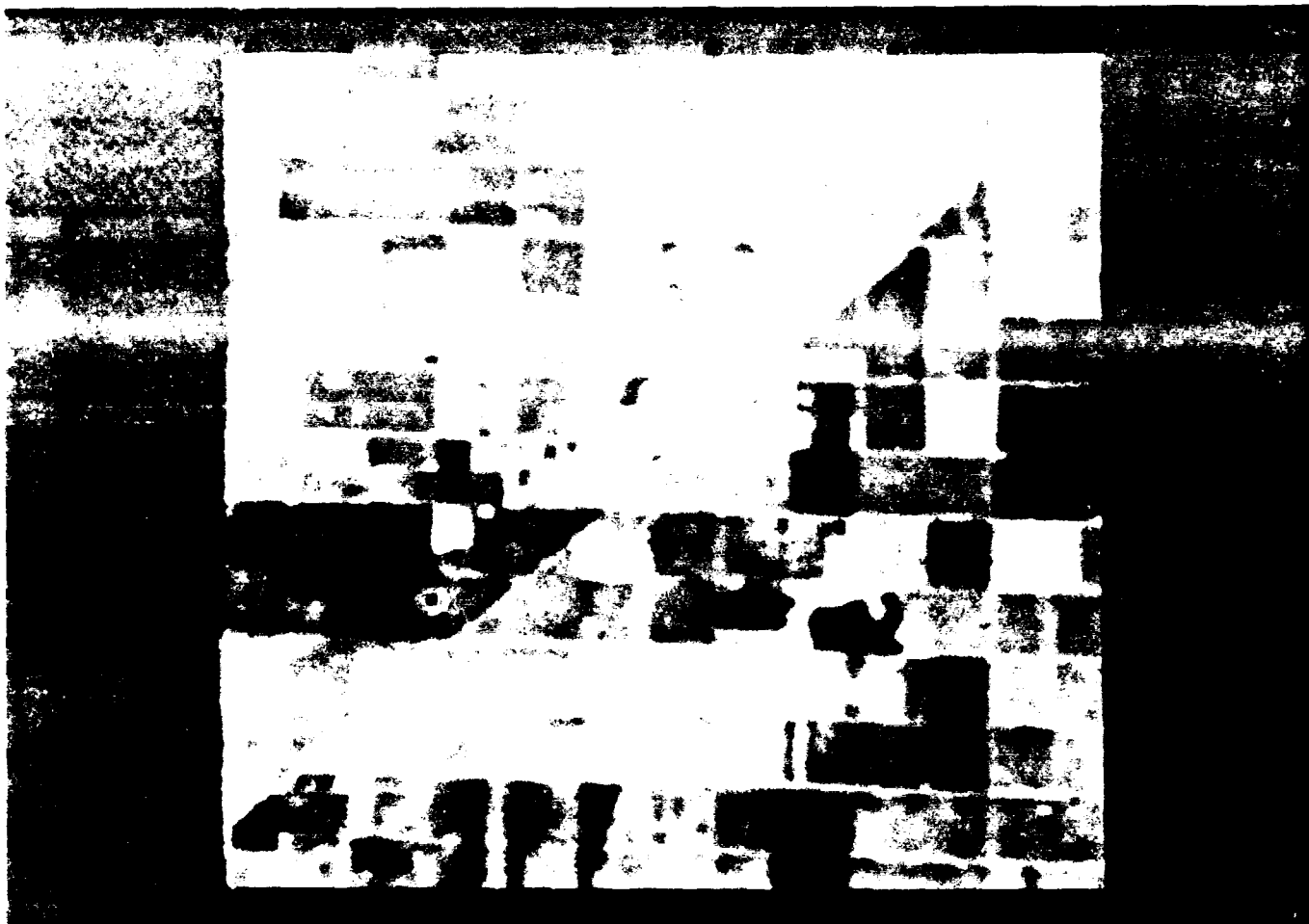


Fig. 1. A. Street in Santa Clara County, Raisin City, CA



Fig. 9-5: Landsat TM Band 5 ($1.55\mu\text{m}$ – $1.75\mu\text{m}$), Raisin City, CA

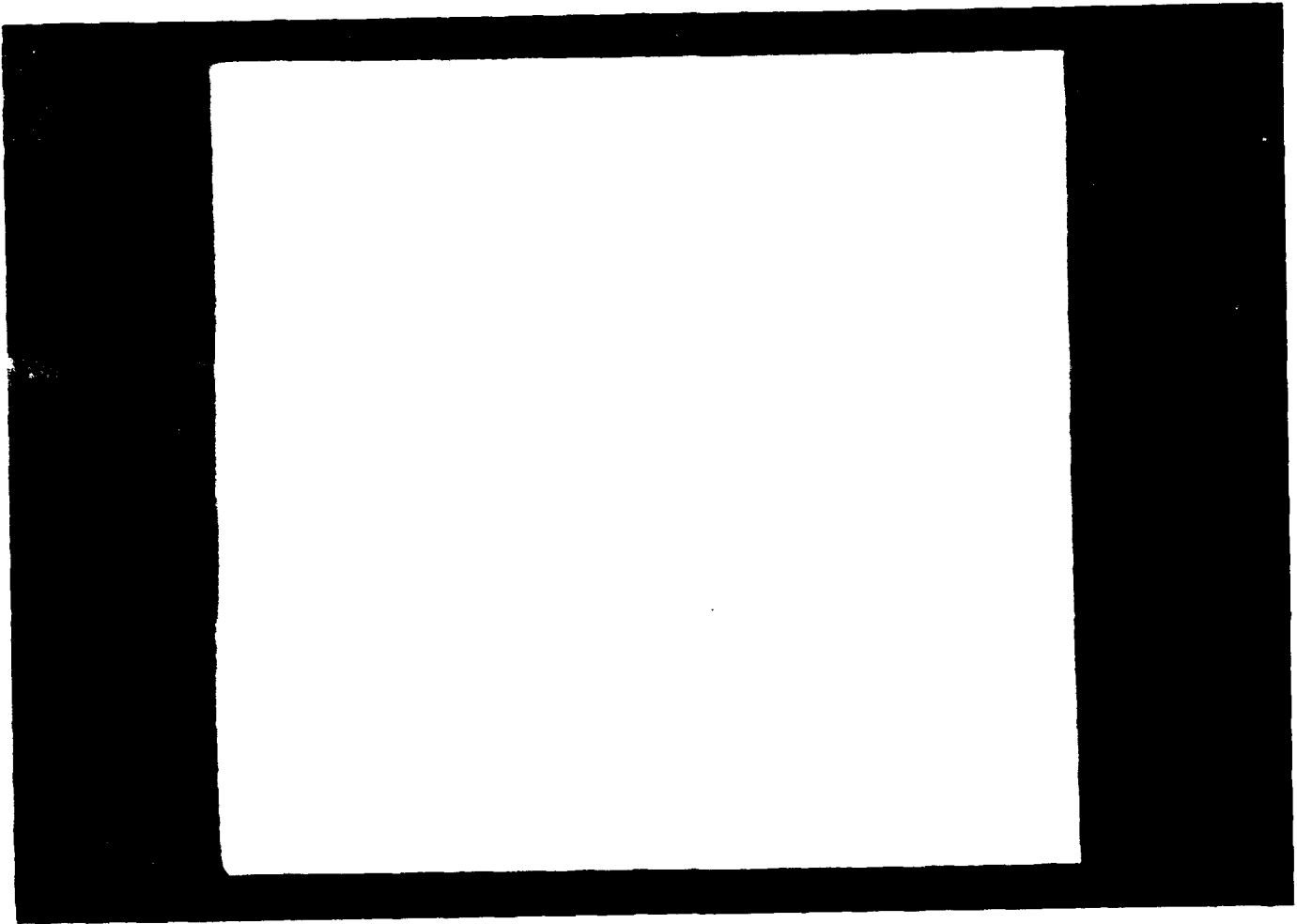


Fig. 9-6: Landsat TM Band 6 (10.4 μ m-.12.5 μ m), Raisin City, CA

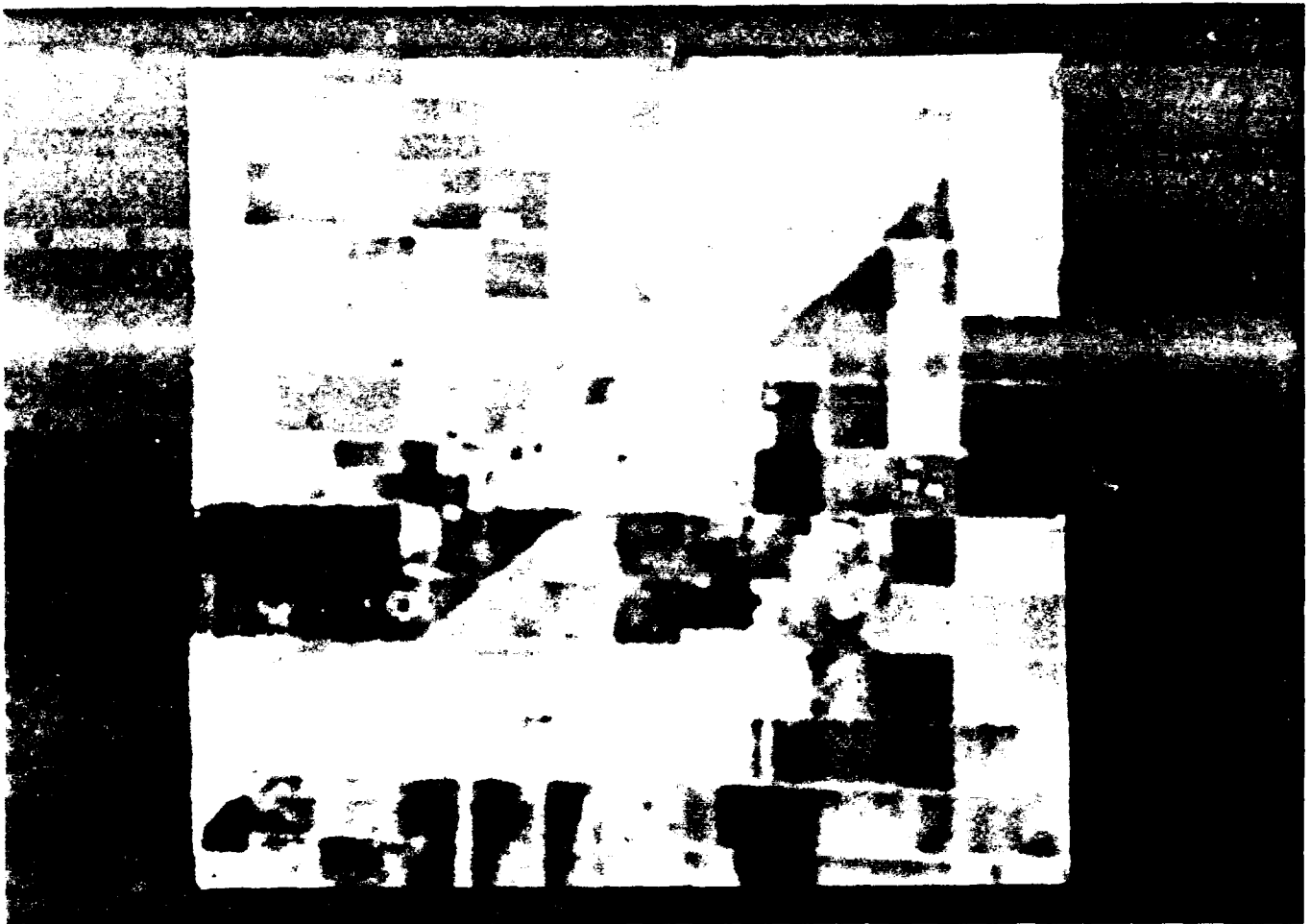


Figure 1. A view of the market area, taken from the street, showing the dense crowd of people and the various stalls and structures.

9.2 Polarimetry Definitions

The following definitions are useful in the analyses of polarimetric measurements:

$\langle \rangle$ = ensemble averaging

$||$ = absolute value of complex quantity

\sim = complex conjugation

$R()$ = real portion of complex quantity (\cdot)

$I()$ = imaginary portion of complex quantity (\cdot)

ψ = polarization ellipse orientation angle, $0 \leq \psi \leq \pi$

χ = polarization ellipse ellipticity angle, $-\pi/4 \leq \chi \leq \pi/4$; $\chi > 0$ corresponds to left-handed sense of polarization, $\chi = 0$ corresponds to linear polarization, and $\chi = \pi/4$ corresponds to full circular polarization

d = fraction of signal power that is fully polarized, $0 \leq d \leq 1$

$\{S\}$ = Stokes vector, a real valued description of polarization

S_0 = first Stokes parameter, corresponds to power

S_1 = second Stokes parameter, $= S_0 d \cos(2\psi) \cos(2\chi)$

S_2 = third Stokes parameter, $= S_0 d \sin(2\psi) \cos(2\chi)$

S_3 = fourth Stokes parameter, $= S_0 d \sin(2\psi)$

$[s]$ = (2x2) complex scattering matrix relating transmitted to received polarization; elements of matrix $[s]$ are as follows:

s_{11} = complex coefficient relating horizontal received component to horizontal transmitted component

s_{12} = complex coefficient relating vertical received component to horizontal transmitted component

s_{21} = complex coefficient relating horizontal received component to vertical transmitted component

s_{22} = complex coefficient relating vertical received component to vertical transmitted component

[F] = (4x4) real Mueller or phase matrix, relating ensemble averaged (time or spatial) received and transmitted polarization as described by Stokes vector; [F] is skew symmetric if [s] is symmetric; elements of matrix [F] are as follows:

$$F_{11} = (|s_{11}|^2 + |s_{12}|^2 + |s_{21}|^2 + |s_{22}|^2)/2$$

$$F_{12} = (|s_{11}|^2 - |s_{12}|^2 + |s_{21}|^2 + |s_{22}|^2)/2$$

$$F_{13} = R(s_{11} * \bar{s}_{12}) + R(s_{22} * \bar{s}_{21})$$

$$F_{14} = I(s_{12} * \bar{s}_{11}) + I(s_{22} * \bar{s}_{21})$$

$$F_{21} = (|s_{11}|^2 + |s_{12}|^2 - |s_{21}|^2 + |s_{22}|^2)/2$$

$$F_{22} = (|s_{11}|^2 - |s_{12}|^2 - |s_{21}|^2 + |s_{22}|^2)/2$$

$$F_{23} = R(s_{11} * \bar{s}_{12}) - R(s_{22} * \bar{s}_{21})$$

$$F_{24} = I(s_{12} * \bar{s}_{11}) + I(s_{21} * \bar{s}_{22})$$

$$F_{31} = R(s_{11} * \bar{s}_{21}) + R(s_{22} * \bar{s}_{12})$$

$$F_{32} = R(s_{11} * \bar{s}_{21}) - R(s_{22} * \bar{s}_{12})$$

$$F_{33} = R(s_{22} * \bar{s}_{11}) + R(s_{12} * \bar{s}_{21})$$

$$F_{34} = I(s_{11} * \bar{s}_{21}) + I(s_{12} * \bar{s}_{22})$$

$$F_{41} = I(s_{11} * \bar{s}_{21}) + I(s_{12} * \bar{s}_{22})$$

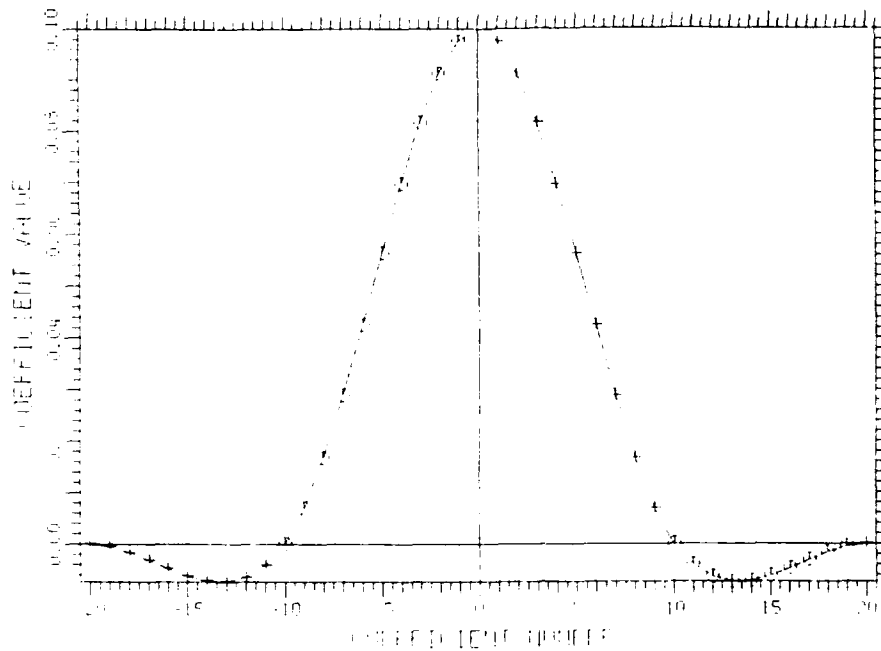
$$F_{42} = I(s_{11} * \bar{s}_{21}) + I(s_{22} * \bar{s}_{12})$$

$$F_{43} = I(s_{11} * \bar{s}_{22}) + I(s_{12} * \bar{s}_{21})$$

$$F_{44} = R(s_{11} * \bar{s}_{22}) - R(s_{12} * \bar{s}_{21})$$

9.3 Resampling Kernels

(a) Impulse Response



(b) Frequency Response

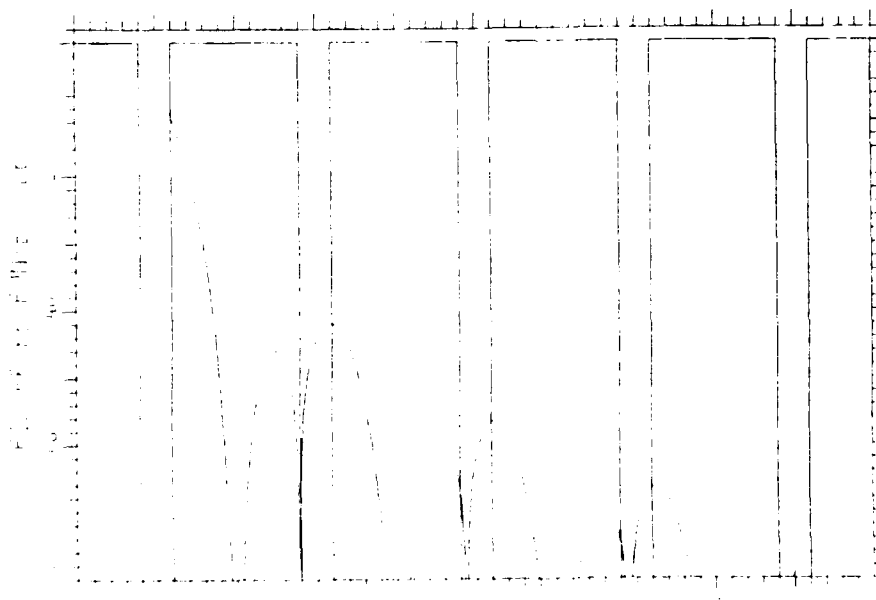
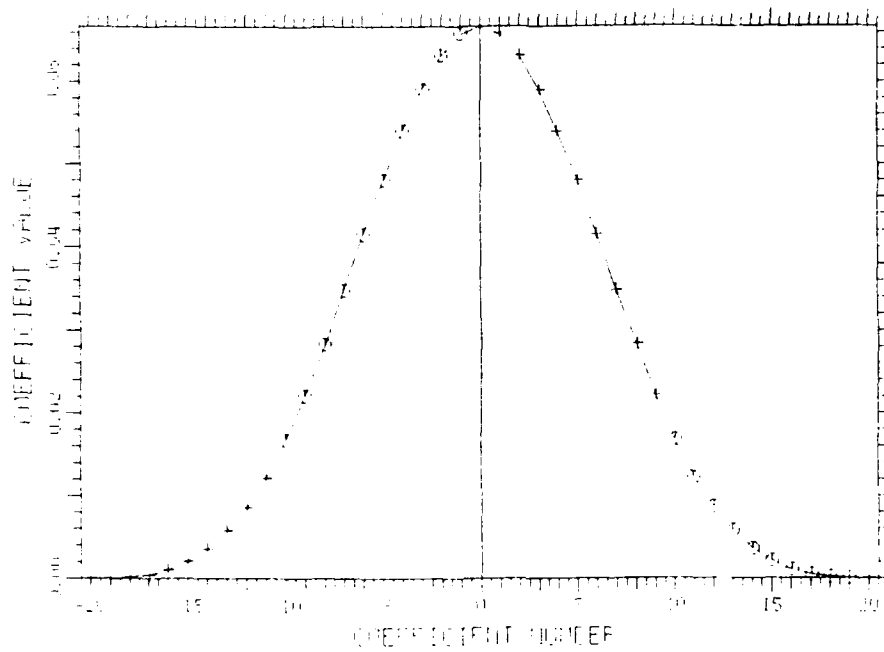


Fig. 9-8: Cubic Convolution Kernel

(a) Impulse Response



(b) Frequency Response

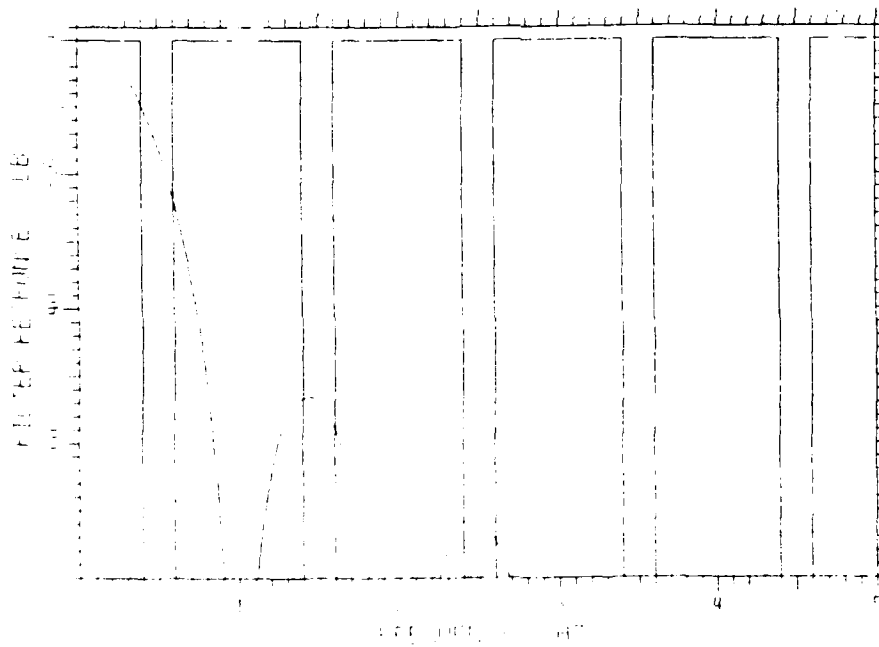
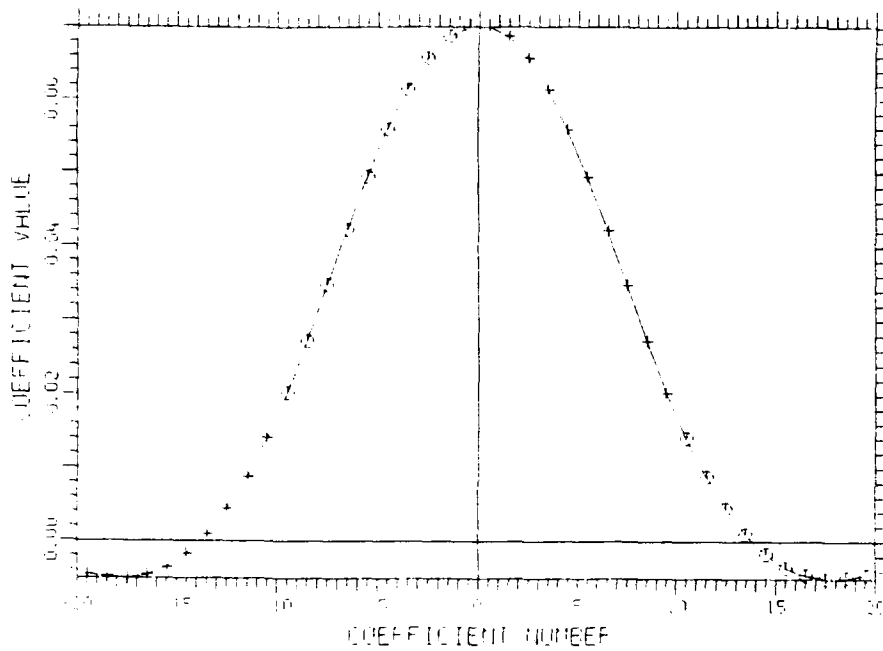


Fig. 9-9: Cubic B-Spline Kernel

(a) Impulse Response



(b) Frequency Response

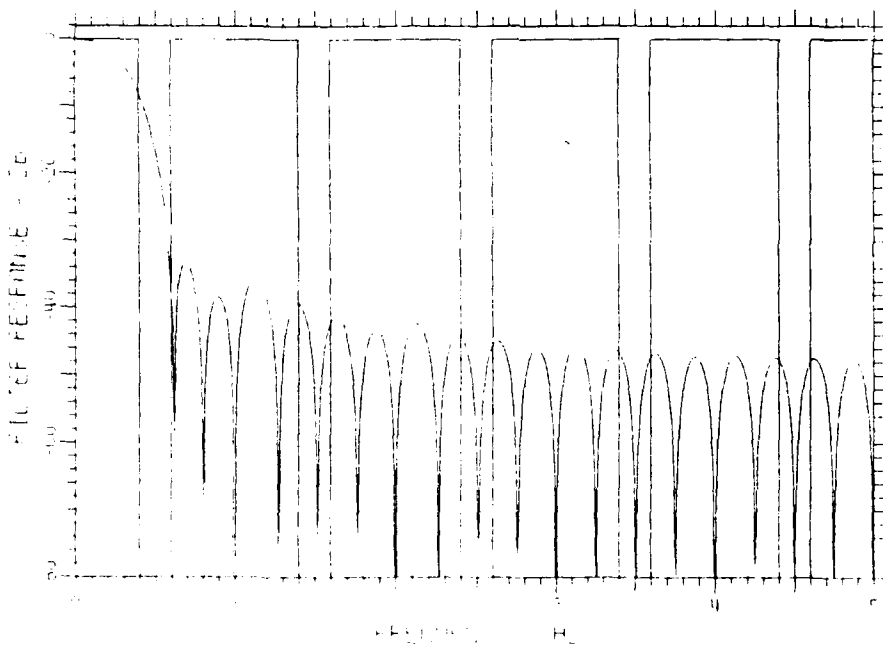
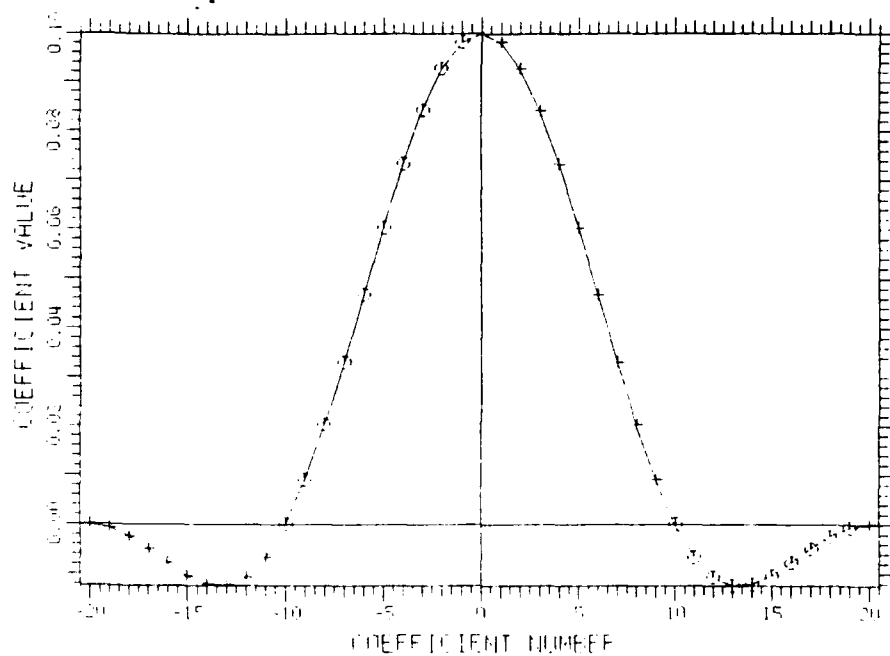


Fig. 9-10: Custom FIR Kernel

(a) Impulse Response



(b) Frequency Response

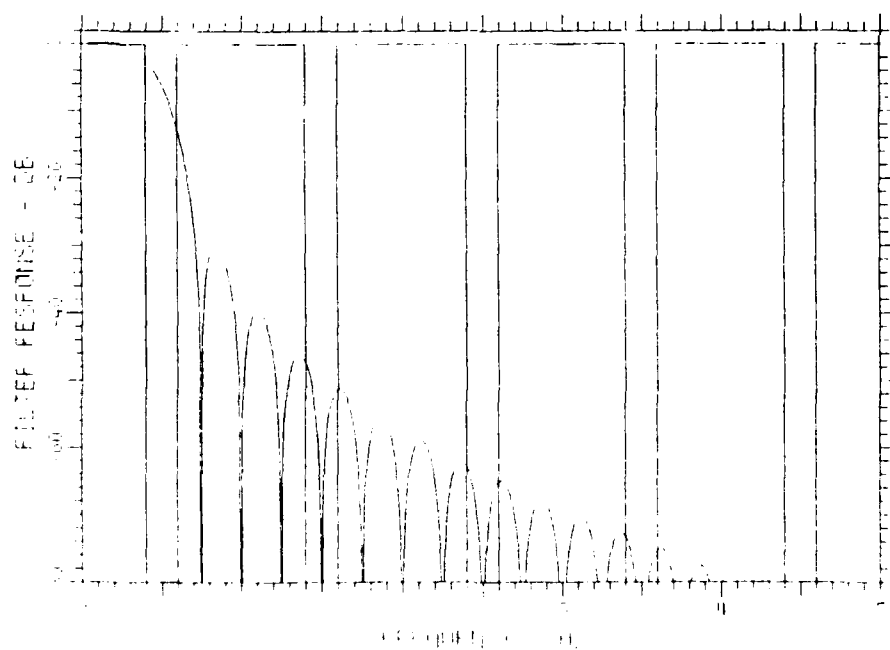
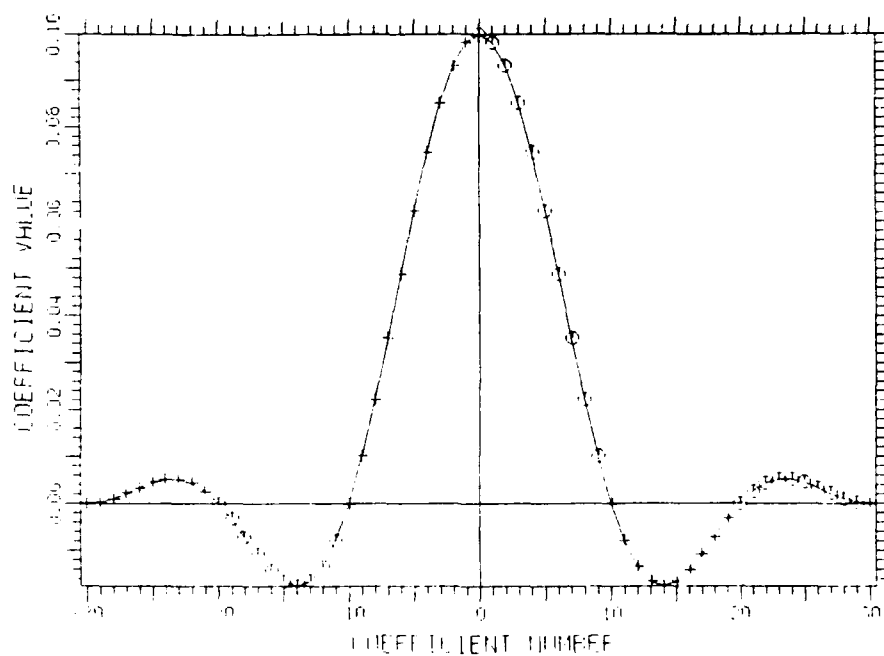


Fig. 9-11: 4-Point DFT Kernel

(a) Impulse Response



(b) Frequency Response

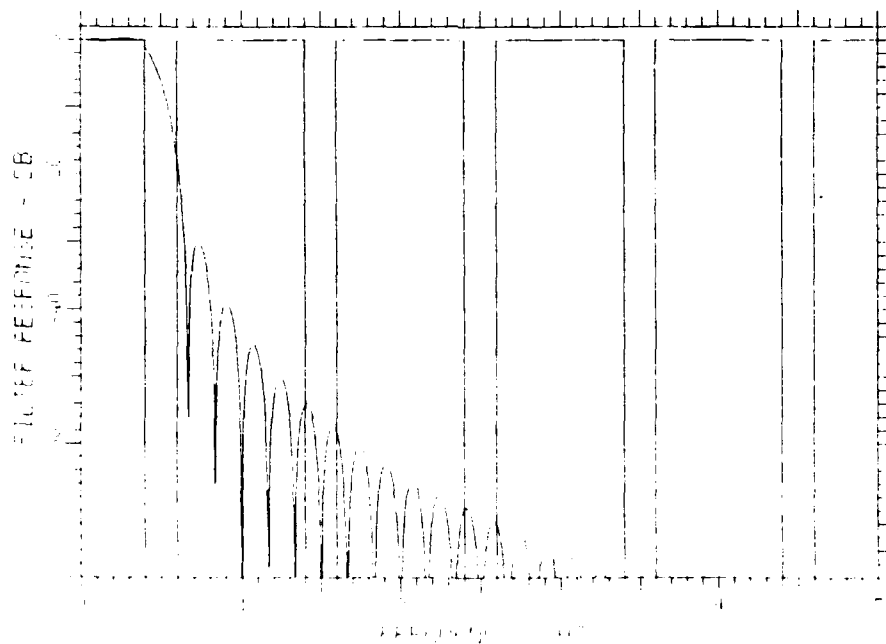
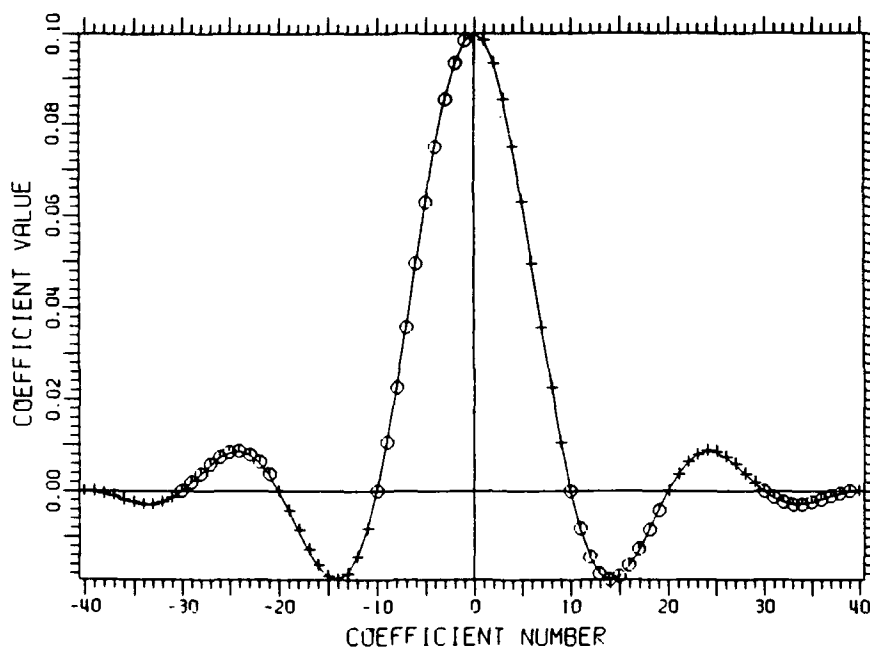


Fig. 9-12: 6-Point DFT Kernel

(a) Impulse Response



(b) Frequency Response

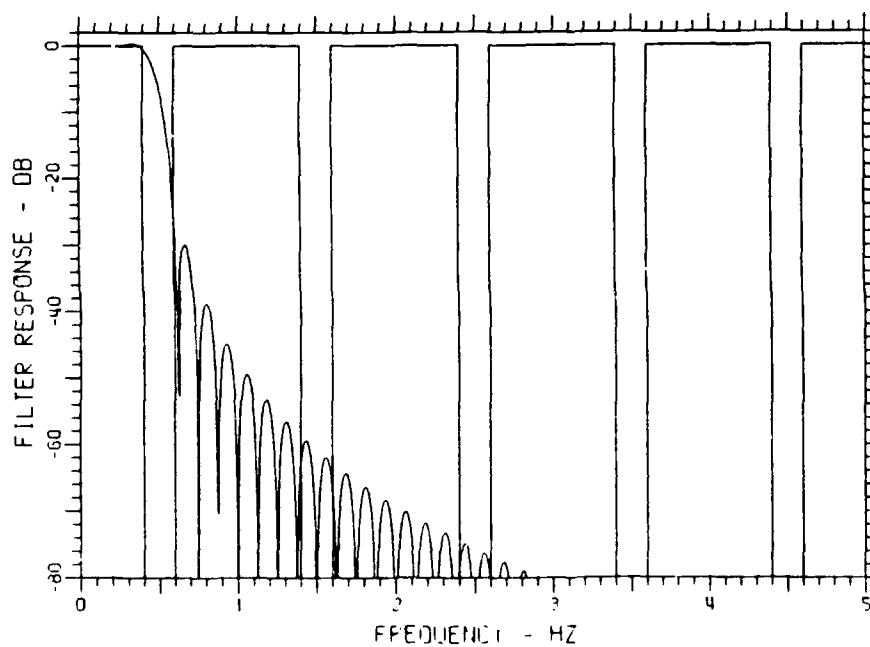


Fig. 9-13: 8-Point DFT Kernel

9.4 LOWTRAN7 Atmospheric Model

LOWTRAN7 is a powerful computer software package developed by the Air Force Geophysics Laboratory to compute the effects of a wide range of atmospheric conditions within a given transmission path. It can calculate the atmospheric transmittance and background radiance, and single and multiple scattered solar and thermal radiance within a spectral resolution of 20 cm^{-1} over the wavelength region $[.2\mu\text{m}, \infty)$.

LOWTRAN uses a single-parameter band model for the molecular absorption and includes molecular scattering, aerosol and hydrometer absorption and scattering. If enough information about a given image sensing scenario is available (time, date, geometry, etc.), then it is possible using LOWTRAN7 to compute the atmospheric contribution to sensed radiance images. Once these effects have been computed, the images can first be converted into source radiance and subsequently to source temperature. The latter step is accomplished by inverting the measured source radiance and the imaging waveband(s).